Difusions (I)

Índex

1. Processos estocàstics

1.1 Processos estocàstics reals indexats en $\mathbb{R}^+$

Un *procés estocàstic real indexat en* $\mathbb{R}^+ = [0, +\infty[$ és una família de variables aleatòries reals $\{X_t,\ t \in \mathbb{R}^+\}$ definides en un mateix espai de probabilitat $(\Omega, \mathfrak{F}, P)$.

Podem pensar-lo com una funció

$$\begin{aligned} \mathbb{R}^+ \times \Omega &\xrightarrow{\quad X \quad} \mathbb{R} \\ (t, \omega) &\longmapsto X_t(\omega) \end{aligned}$$

De la mesurabilitat de cada X_t no es dedueix que aquesta funció sigui mesurable respecte la σ-àlgebra producte $\mathfrak{B}(\mathbb{R}^+) \otimes \mathfrak{F}$. Quan ho és, diem que el procés és *mesurable.*

Per a cada $\omega \in \Omega$, el procés defineix una funció

$$\begin{aligned} X_\cdot(\omega) \colon \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ t &\longmapsto X_t(\omega) \end{aligned}$$

que s'anomena *trajectòria del procés.* Per tant, també podem veure X com una aplicació

$$\begin{aligned} X \colon \Omega &\longrightarrow \mathbb{R}^{\mathbb{R}^+} \\ \omega &\longmapsto X_\cdot(\omega) \end{aligned} \tag{1.1.1}$$

on $\mathbb{R}^{\mathbb{R}^+}$ és una notació pel producte cartesià $\underset{t \in \mathbb{R}^+}{\times} \mathbb{R}$ (de "$\mathbb{R}^+$ còpies" de $\mathbb{R}$), que no és més que el conjunt de totes les funcions $\mathbb{R}^+ \to \mathbb{R}$.

1.2 Llei d'un procés estocàstic

Donat un espai de mesura $(\Omega, \mathfrak{F}, \mu)$ i una funció mesurable X a valors en un altre espai de mesura $(E, \mathfrak{E})$, es defineix la *mesura imatge* de μ per X com la mesura μ_X en $(E, \mathfrak{E})$ definida per

$$\mu_X(B) = \mu(X^{-1}(B)) \quad , \quad B \in \mathfrak{E} \quad .$$

Quan μ és una probabilitat, X s'anomena *variable aleatòria E-valuada* i a la mesura imatge corresponent l'anomenem la *llei de X*.

Si pensem un procés estocàstic real com l'aplicació (1.1.1) i posem en $\mathbb{R}^{\mathbb{R}^+}$ una σ-àlgebra tal que X sigui mesurable, tindrem de fet una variable aleatòria $\mathbb{R}^{\mathbb{R}^+}$-valuada, i podrem parlar de la *llei del procés X*.

Quina σ-àlgebra posem en $\mathbb{R}^{\mathbb{R}^+}$? La σ-àlgebra natural en un producte cartesià és la *σ-àlgebra producte*, que es defineix com la més petita que fa mesurables les projeccions

$$\begin{aligned} \pi_t\colon \mathbb{R}^{\mathbb{R}^+} &\longrightarrow \mathbb{R} \\ f &\longmapsto f(t); \end{aligned}$$

equivalentment, com la generada pels conjunts

$$\{\pi_t^{-1}(B) :\ B \in \mathfrak{B}(\mathbb{R}),\ t \in \mathbb{R}^+\} \quad ;$$

equivalentment, com la generada pels *cilindres:*

$$\{\pi_{t_1,\dots,t_n}^{-1}(B_1 \times \cdots \times B_n) :\ B_i \in \mathfrak{B}(\mathbb{R}),\ t_i \in \mathbb{R}^+,\ n \in \mathbb{N}\} \quad ;$$

i de fet és exactament la σ-àlgebra formada pels *σ-cilindres:*

$$\{\pi_J^{-1}(B_J) :\ B_J \in \mathfrak{B}(\mathbb{R})^J,\ J \subset \mathbb{R}^+ \text{ numerable}\} \quad ,$$

on $\mathfrak{B}(\mathbb{R})^J$ és el producte de "card J còpies" de $\mathfrak{B}(\mathbb{R})$, i π_J és la projecció natural de $\mathbb{R}^{\mathbb{R}^+}$ en $\mathbb{R}^J$. Notem per $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$ aquesta σ-àlgebra producte.

En el paràgraf anterior, estem parlant de l'espai de mesura producte de còpies de $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. Per la situació general (espai mesurable producte d'una família d'espais mesurables $\{(\Omega_i, \mathfrak{F}_i)\}_{i\in I}$), vegeu Williams [13], II.5 (tot funciona igual).

Es comprova que si $\{X_t :\ t \in \mathbb{R}^+\}$ és un procés estocàstic (és a dir, cada $X_t\colon \Omega \to \mathbb{R}$ és mesurable), i considerem en $\mathbb{R}^{\mathbb{R}^+}$ la σ-àlgebra producte, aleshores $X\colon \Omega \to \mathbb{R}^{\mathbb{R}^+}$ és mesurable.

Seguim parlant de l'espai mesurable producte $(\mathbb{R}^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+})$: Donada una probabilitat P en aquest espai, podem considerar, per a cada $J \subset \mathbb{R}^+$ finit, la mesura imatge de P per la projecció π_J, definida com

$$P_J \equiv P \circ \pi_J^{-1} \quad . \tag{1.2.1}$$

Aquesta P_J és una probabilitat sobre un producte finit de còpies de $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. La família de probabilitats $\{P_J :\ J \subset \mathbb{R}^+ \text{ finit}\}$ s'anomena *distribucions en dimensió finita de P*, i compleixen, si $K \subset J$,

$$P_K \equiv P_J \circ \pi_{J,K}^{-1} \quad , \tag{1.2.2}$$

on $\pi_{J,K}$ és la projecció natural de $\mathbb{R}^J$ en $\mathbb{R}^K$.

Suposem que tenim ara una família de probabilitats $\{P_J,\ J \subset \mathbb{R}^+ \text{ finit}\}$, cadascuna sobre $\mathbb{R}^J$, complint (1.2.2), anomenada *propietat de consistència.* La col·lecció dels cilindres de $\mathbb{R}^{\mathbb{R}^+}$ és una àlgebra que genera la σ-àlgebra producte. La propietat de consistència ens assegura que es pot definir una funció de conjunt additiva P sobre els cilindres que compleix (1.2.1). Els teoremes de Carathéodory ens diuen que P es pot estendre a una probabilitat (única) sobre la σ-àlgebra producte si i només si P és numerablement additiva sobre els cilindres. L'extensió s'anomena la *probabilitat límit projectiu* de la família $\{P_J :\ J \subset \mathbb{R}^+ \text{ finit}\}$, i compleix (1.2.1).

I, en efecte, aquesta additivitat numerable és certa. És un cas particular del Teorema de Daniell–Kolmogorov (Teorema 1.5.1).

De la discussió anterior, obtenim que la llei d'un procés estocàstic queda determinada per les *lleis en dimensió finita:* $\{P_{X_{t_1},\dots,X_{t_n}} : t_1 < \cdots < t_n \in \mathbb{R}^+\}$.

1.2.1 Observació

Es pot demostrar que $\mathfrak{B}(\mathbb{R}^n)$ (considerant en $\mathbb{R}^n$ la topologia producte, que és l'habitual de $\mathbb{R}^n$) coincideix amb la σ-àlgebra producte de n còpies de $\mathfrak{B}(\mathbb{R})$, o sigui, que $\mathfrak{B}(\mathbb{R}^n) = \mathfrak{B}(\mathbb{R})^n$. El mateix és cert amb una potència numerable, és a dir, $\mathfrak{B}(\mathbb{R}^{\mathbb{N}}) = \mathfrak{B}(\mathbb{R})^{\mathbb{N}}$.
En canvi, no ho és per una potència no numerable: La σ-àlgebra producte d'una quantitat no numerable de còpies de $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ és més petita que la σ-àlgebra dels borelians de l'espai topològic producte. Vegeu en el Corol·lari 2.4.2 una situació en la que cal anar amb compte amb això.
Hi ha una explicació molt bona de perquè passa això en els paràgrafs (4) a (8) de Kendall [6]. Destaquem que en aquest fenomen la cardinalitat de $\mathbb{R}$ no n'és responsable: el mateix passa amb conjunts de dos punts.

1.3 Representació canònica

Considerem la pregunta elemental següent: Donada una probabilitat P en $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, existeix alguna variable aleatòria que la tingui per llei? La resposta és sí, trivialment: Prenem $(\Omega, \mathfrak{F}) = (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, hi posem com a probabilitat la mateixa P, i definim $X : \mathbb{R} \to \mathbb{R}$ com $X(\omega) = \omega$ (la identitat). Podem dir que X és la variable *canònica* corresponent a aquesta llei.

El raonament funciona exactament igual substituint $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), P)$ per un espai de probabilitat qualsevol. Per tant, sempre existeix un procés estocàstic amb una llei donada. La variable aleatòria $X : \Omega \to \mathbb{R}^{\mathbb{R}^+}$ és la identitat, i se l'anomena procés canònic relatiu a la llei P.

Observació tonta: En el planteig anterior tenim llibertat per escollir l'espai mesurable de base $(\Omega, \mathfrak{F})$. Si aquest no fos el cas, la resposta podria ser negativa; per exemple, si Ω és finit, no es pot tenir definida en $(\Omega, \mathfrak{F})$ una variable aleatòria amb llei $N(0, 1)$. L'espai $(\Omega, \mathfrak{F})$ ha de ser "prou ric", per tal que es pugui fer.

Al prendre $(\Omega, \mathfrak{F}, P)$ el propi espai de probabilitat on tenim definida la pressumpta llei, estem fent també una elecció "canònica"; és el *espai canònic* relatiu a aquesta llei.

Donat un procés $Y : \Omega \to \mathbb{R}^{\mathbb{R}^+}$, sobre un espai de probabilitat qualsevol $(\Omega, \mathfrak{F}, P)$, amb llei $P_Y = P \circ Y^{-1}$, el *procés canònic* associat amb Y és el procés definit per

$$X_t(\omega) = \omega(t)$$

en l'espai de probabilitat

$$(\mathbb{R}^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}, P_Y) \quad .$$

Estudiar processos estocàstics es redueix a estudiar la variable identitat? No. Segons el cas, es convenient o no utilitzar la representació canònica d'un procés. Amb la representació canònica, les trajectòries són senzilles de representar, però les lleis de probabilitat poden ser complicades. Pot convenir tenir la situació contrària. Per altra banda, sempre que hàgim de considerar a la vegada dos processos X, Y, no podem òbviament prendre la representació canònica per tots dos a la vegada. Igual passa amb variables aleatòries valuades a $\mathbb{R}$; per exemple, X i $-X$ no poden ser la identitat totes dues a la vegada, en general.

1.4 Extensions de la representació canònica

Suposem que volem parlar d'una certa propietat de les trajectòries d'un procés estocàstic; per exemple, volem dir que un cert procés té trajectòries contínues amb probabilitat 1. Això equival a dir que $P_X(C(\mathbb{R}^+)) = 1$.

Moltes de les "propietats bones" que podem demanar a les trajectòries d'un procés són d'aquest tipus: Hi ha un cert conjunt $\Gamma \subset \mathbb{R}^{\mathbb{R}^+}$ i la propietat s'expressa com "$X(\omega) \in \Gamma$".

Però pot passar fàcilment que $\Gamma \notin \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$. L'espai $\mathbb{R}^{\mathbb{R}^+}$ és molt gran i la σ-àlgebra $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$ és molt petita. Com a regla bastant fiable a seguir, si la propietat en qüestió depèn només d'una quantitat numerable de coordenades de la trajectòria, serà a $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$; si depèn de manera essencial d'una quantitat no numerable, no hi serà. Així, per exemple, $C(\mathbb{R}^+) \notin \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$.

Vejam com se solucionen aquestes problemes:

Suposem que tenim un espai de probabilitat $(\Omega, \mathfrak{F}, P)$, i sigui $Y \equiv \{Y_t,\ t \in \mathbb{R}^+\}$ un procés estocàstic sobre aquest espai.

Siguin $(\mathbb{R}^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}, P_Y)$ l'espai canònic i $X_t(\omega) = \omega(t)$ el procés canònic.

Suposem que Y té trajectòries amb les propietats $\{\Gamma_\lambda\}_{\lambda\in\Lambda}$, en el sentit precís següent:

Per a cada λ, $Y_\cdot(\omega) \in \Gamma_\lambda \subset \mathbb{R}^{\mathbb{R}^+}$, per a tot ω fora d'un conjunt $N_\lambda \in \mathfrak{F}$, amb $P(N_\lambda) = 0$.

Podem suposar sense perdre generalitat que tota intersecció numerable de conjunts de $\{\Gamma_\lambda\}_{\lambda\in\Lambda}$ és altre cop de $\{\Gamma_\lambda\}_{\lambda\in\Lambda}$ (si no hi és, la hi posem).

Definim l'aplicació

$$\begin{array}{ccc} (\Omega, \mathfrak{F}, P) & \xrightarrow{\quad T \quad} & (\mathbb{R}^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}, P_Y) \\ \omega & \longmapsto & Y_\cdot(\omega) \end{array}$$

que assigna a cada $\omega \in \Omega$ la trajectòria $Y_\cdot(\omega)$. És trivial que:

a) T és mesurable.

b) La mesura imatge de P per T és P_Y.

c) $Y = X \circ T$.

Ara, suposem primer que $\Gamma_\lambda \in \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$. Aleshores $T^{-1}(\Gamma_\lambda)$ és un conjunt de $\mathfrak{F}$ que conté $\Omega - N_\lambda$. Com que $P(\Omega - N_\lambda) = 1$, tindrem $P_Y(\Gamma_\lambda) = 1$, i no hi ha cap problema.

Si $\Gamma_\lambda \notin \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$, encara tindrem que, si B és tal que $\Gamma_\lambda \subset B \in \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$,

$$T^{-1}(B) \supset T^{-1}(\Gamma_\lambda) \supset \Omega - N_\lambda$$

i $P_Y(B) = P(T^{-1}(B)) = 1$, com abans. Això implica que Γ_λ té mesura exterior 1 respecte $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$ i P_Y.

En resum, tenim que, $\forall\lambda \in \Lambda$, Γ_λ té $(\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}, P_Y)$-mesura exterior 1.

Sigui $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}[\Gamma]$ la més petita σ-àlgebra que conté $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$ i tots els conjunts Γ_λ. Resulta que existeix una única extensió de P_Y a $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}[\Gamma]$ que dóna mesura 1 a tots els Γ_λ. És intuïtiu, però no evident. Per una demostració, vegeu Meyer [9], Capítol II, Teorema 27.

Amb tot això, arribem a l'enunciat següent:

1.4.1 Teorema

Donats un procés Y definit en un espai qualsevol $(\Omega, \mathfrak{F}, P)$, i una família $\{\Gamma_\lambda\}_{\lambda\in\Lambda}$, tancada per intersecció numerable, de "bones propietats" que les trajectòries de Y poseeixen P-q.s., existeix una única extensió minimal

$$(\mathbb{R}^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}[\Gamma], P_Y[\Gamma])$$

de l'espai canònic

$$(\mathbb{R}^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}, P_Y)$$

que fa mesurable cada Γ_λ i li dóna probabilitat 1.

1.4.2 Observació

La probabilitat a l'espai estès depèn també de la família $\{\Gamma_\lambda\}_{\lambda\in\Lambda}$, perquè dues famílies diferents podrien ambdues fer mesurable un cert conjunt Γ de mesura exterior 1 respecte $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$ i P_Y, i les extensions donar-li mesures diferents.
Això pot passar perquè poden existir conjunts amb mesura exterior 1 tals que el seu complementari també té mesura exterior 1 (vegeu l'exemple que segueix). Un tal conjunt tindrà mesura 1 en una extensió i mesura zero en una altra.

1.4.3 Exemple

Sigui $\Gamma \subset \mathbb{R}^{\mathbb{R}^+}$ el conjunt format només per la funció idènticament zero. Γ és un borelià de $\mathbb{R}^{\mathbb{R}^+}$, però $\Gamma \notin \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$.
Considerem la probabilitat μ sobre els borelians que fa $\mu(\Gamma) = 1$. La restringim a $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$. Aleshores Γ tindrà mesura exterior 1, però Γ^c també, perquè l'únic element de $\mathfrak{B}(\mathbb{R})^{\mathbb{R}^+}$ que el conté és el total $\mathbb{R}^{\mathbb{R}^+}$. $\square$

Una conclusió d'aquest exemple és que no hi pot haver una extensió canònica universal, que estengui totes les altres. Però el que importa és que sempre podrem "realitzar" les bones propietats que tingui un procés definit en un espai qualsevol mitjançant una extensió adequada de l'espai canònic.

1.5 Processos estocàstics més generals

Parlarem sempre de processos estocàstics a temps continu amb espai de paràmetres $\mathbb{R}^+ = [0, +\infty[$, llevat que s'indiqui el contrari. Una situació més general seria $T \subset \mathbb{R}$, i per certes questions no seria indiferent si T té mínim o no. La situació més general correspon a considerar com a espai de paràmetres un conjunt qualsevol T.

Quant a *l'espai d'estats* (l'espai mesurable on prenen els seus valors les variables del procés), serà habitualment $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ (processos reals) o $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$. La situació més general és aquella en què l'espai d'estats és un espai mesurable qualsevol.

El Teorema de Daniell–Kolmogorov ens assegura l'existència de processos amb una família consistent fixada de distribucions en dimensió finita en una situació força general.

1.5.1 Teorema de Daniell–Kolmogorov (Daniell,1918; Kolmogorov, 1933)

Sigui E un espai polonès (i.e. mètric, separable i complet), i considerem l'espai mesurable $(E, \mathfrak{B}(E))$.
Sigui I un conjunt d'índexos de cardinalitat arbitrària.
Sigui $\{P_J,\ J \subset I \text{ finit}\}$ una família de probabilitats complint la propietat de consistència (1.2.2).

Aleshores existeix la probabilitat límit projectiu de la família sobre l'espai mesurable $(E^I, \mathfrak{B}(E)^I)$. □

El fet crucial que permet que l'enunciat funcioni és que tota probabilitat P sobre un espai polonès (sempre amb la σ-àlgebra de Borel) és *ajustada:* Per a tot B borelià, existeix un compacte $K \subset B$ tal que

$$P(B - K) < \varepsilon \quad .$$

El Teorema de Daniell–Kolmogorov ens dóna immediatament l'existència de processos amb una família de distribucions en dimensió finita donada, en el cas E polonès (almenys tindrem segur el procés canònic).

Observació marginal: Ocasionalment, es pot veure el Teorema 1.5.1 amb les hipòtesis: E Hausdorff, compacte i complint el segon axioma de numerabilitat (i.e. la topologia té una base numerable). Però resulta que aquestes hipòtesis impliquen (no és del tot trivial) que l'espai és separable, metritzable i tota mètrica dóna lloc necessàriament a un espai mètric complet. Per tant, això no és important, i l'enunciat 1.5.1 és més general.

2. Esperança i probabilitat condicionades

Sobre esperança i probabilitat condicionades, hi ha moltes idees intuïtives a Ash [2], Cap. 6, i Williams [14], Cap. 9.

2.1 Esperança condicionada

Una *mesura signada* (o *mesura real*) μ en un espai mesurable $(\Omega, \mathfrak{F})$ és una funció de conjunt $\mu: \mathfrak{F} \to \mathbb{R}$ σ-additiva. El Teorema de Radon-Nikodým és valid per a mesures reals, en el sentit següent: Si μ és una mesura signada i ν és una mesura positiva σ-finita, aleshores μ és absolutament contínua respecte ν (i.e. $\nu(A) = 0 \Rightarrow \mu(A) = 0$) sii $\exists f \in L^1(\Omega, \mathfrak{F}, \nu)$ tal que $\mu(A) = \int_A f \, d\nu$. Aquesta f és única ν-q.p.t.

2.1.1 Proposició

Sigui $(\Omega, \mathfrak{F}, P)$ *un espai de probabilitat.*
Sigui $\mathfrak{G}$ *una sub-*σ*-àlgebra de* $\mathfrak{F}$*.*
Sigui $X \in L^1(\Omega, \mathfrak{F}, P)$*.*
Aleshores, existeix un únic element $\mathrm{E}\left[X/\mathfrak{G}\right] \in L^1(\Omega, \mathfrak{G}, P)$ *tal que*

$$\int_B X \, dP = \int_B \mathrm{E}\left[X/\mathfrak{G}\right] dP \quad , \quad \forall B \in \mathfrak{G} \quad .$$

Demostració: Sigui μ la mesura signada definida per

$$\mu(A) = \int_A X \, dP \quad , \quad A \in \mathfrak{F} \quad .$$

μ és absolutament contínua respecte P. Ho segueix sent si considerem μ i P definides només sobre $\mathfrak{G}$. Pel Teorema de Radon–Nikodým, existirà una densitat (signada) f, mesurable respecte $\mathfrak{G}$, tal que

$$\mu(B) = \int_B f \, dP \quad , \quad B \in \mathfrak{G} \quad .$$

$\mathrm{E}\left[{}^{X}\!/_{\mathfrak{G}}\right] := f$ compleix òbviament les propietats enunciades, i és única com a element de $L^1(\Omega, \mathfrak{G}, P)$.

2.1.2 Definició

La variable aleatòria $\mathrm{E}\left[{}^{X}\!/_{\mathfrak{G}}\right]$ de la Proposició 2.1.1 s'anomena *esperança de X condicionada a $\mathfrak{G}$.*

2.1.3 Exemple

Si $\mathfrak{G}$ està generada per una partició $B_1, \dots, B_n$ de Ω, amb $P(B_i) > 0$, $\forall i$, aleshores

$$\mathrm{E}\left[{}^{X}\!/_{\mathfrak{G}}\right] = \sum_{i=1}^{n} \Big(\frac{1}{P(B_i)} \int_{B_i} X\,dP\Big) \cdot \mathbf{1}_{B_i} \quad .$$

Més en general, si B és un *àtom* de $\mathfrak{G}$ (i.e. $A \in \mathfrak{G}$, $A \subset B \Rightarrow P(A) = 0$ o $P(B - A) = 0$) amb $P(B) > 0$, aleshores

$$\mathrm{E}\left[{}^{X}\!/_{\mathfrak{G}}\right] = \frac{1}{P(B)} \int_B X\,dP \quad , \quad \text{q.s. sobre } B \quad .$$

2.2 Definició alternativa de l'esperança condicionada

La següent és una construcció alternativa de la esperança condicionada, força elegant i que en proporciona una caracterització útil.

Considerem l'espai de Hilbert $L^2(\Omega, \mathfrak{F}, P)$. Sigui $\mathfrak{G}$ una sub-σ-àlgebra de $\mathfrak{F}$. L'espai $L^2(\Omega, \mathfrak{G}, P)$ és un subspai tancat de $L^2(\Omega, \mathfrak{F}, P)$.

2.2.1 Definició

Es defineix *l'operador esperança condicionada a* $\mathfrak{G}$ en $L^2(\Omega, \mathfrak{F}, P)$ com la projecció sobre $L^2(\Omega, \mathfrak{G}, P)$:

$$\begin{aligned} \pi\colon L^2(\Omega, \mathfrak{F}, P) &\longrightarrow L^2(\Omega, \mathfrak{G}, P) \\ \xi &\longmapsto \pi(\xi) =: \mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right] \end{aligned}$$

2.2.2 Proposició

$\mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right]$ *és l'únic element de* $L^2(\Omega, \mathfrak{G}, P)$ *tal que*

$$\int_B \xi\,dP = \int_B \mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right] dP \quad , \quad \forall B \in \mathfrak{G} \quad ,$$

i és per tant l'esperança condicionada definida abans.

Demostració: El conjunt de funcions $\{\mathbf{1}_B,\ B \in \mathfrak{G}\}$ formen un sistema total en $L^2(\Omega, \mathfrak{G}, P)$. Per tant, $\mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right]$ queda únicament determinat pel fet que $\xi - \mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right] \perp \mathbf{1}_B$, $\forall B \in \mathfrak{G}$. O sigui,

$$\int_\Omega \big(\xi - \mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right]\big) \cdot \mathbf{1}_B\,dP = 0 \quad ,$$

d'on surt immediatament l'enunciat. $\square$

Finalment, $L^2(\Omega, \mathfrak{F}, P)$ és un subspai dens de $L^1(\Omega, \mathfrak{F}, P)$. Si comprovem que l'operador esperança condicionada és acotat en la norma de L^1, s'entendrà per continuïtat a un operador acotat en $L^1(\Omega, \mathfrak{F}, P)$, i queda completa la identificació amb la construcció de l'Apartat 2.1. En efecte,

$$\|\,\mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right]\|_{L^1} = \mathrm{E}\left[|\,\mathrm{E}\left[{}^{\xi}\!/_{\mathfrak{G}}\right]|\right] \leqslant \mathrm{E}\left[\,\mathrm{E}\left[{}^{|\xi|}\!/_{\mathfrak{G}}\right]\right] = \mathrm{E}[|\xi|] = \|\xi\|_{L^1} \quad ,$$

aplicant la Desigualtat de Jensen, la Propietat de la Torre, i la Conservació de l'Esperança de l'apartat següent.

2.3 Propietats de l'esperança condicionada

Això és una llista de propietats de l'esperança condicionada. Es poden demostrar només per a variables de L^2, i s'estenen automàticament a variables de L^1, per continuïtat, segons l'extensió de l'Apartat 2.2. Per les demostracions, vegeu, per exemple, Williams [14]. Se suposa que totes les variables aleatòries que hi apareixen són de L^1.

1) *Conservació de l'esperança:* $\mathrm{E}\left[\mathrm{E}\left[X/_{\mathfrak{G}}\right]\right]=\mathrm{E}[X]$.
2) *Linealitat:* $\mathrm{E}\left[a_1X_1+a_2X_2/_{\mathfrak{G}}\right]=a_1\,\mathrm{E}\left[X_1/_{\mathfrak{G}}\right]+a_2\,\mathrm{E}\left[X_2/_{\mathfrak{G}}\right]$.
3) *Positivitat:* $X\geqslant 0\Rightarrow \mathrm{E}\left[X/_{\mathfrak{G}}\right]\geqslant 0$. Corol·lari: *Monotonia*.
4) *Convergència monòtona:* $0<X_n\nearrow X$, $\forall n\in\mathbb{N}$, q.s., $\Rightarrow \mathrm{E}\left[X_n/_{\mathfrak{G}}\right]\nearrow \mathrm{E}\left[X/_{\mathfrak{G}}\right]$.
5) *Fatou:* $0<X_n$, $\forall n\in\mathbb{N}$, q.s., $\Rightarrow \mathrm{E}\left[\varliminf_{n\to\infty} X_n/_{\mathfrak{G}}\right]\leqslant \varliminf_{n\to\infty}\mathrm{E}\left[X_n/_{\mathfrak{G}}\right]$.
6) *Convergència dominada:* $|X_n|\leqslant Y\in L^1$, $\forall n\in\mathbb{N}$, i $X_n\to X$, q.s., $\Rightarrow \mathrm{E}\left[X_n/_{\mathfrak{G}}\right]\to \mathrm{E}\left[X/_{\mathfrak{G}}\right]$.
7) *Desigualtat de Jensen:* $\varphi\colon\mathbb{R}\to\mathbb{R}$ convexa, $\varphi(X)\in L^1$, $\Rightarrow \varphi\big(\mathrm{E}\left[X/_{\mathfrak{G}}\right]\big)\leqslant \mathrm{E}\left[\varphi(X)/_{\mathfrak{G}}\right]$. En particular, $\left|\mathrm{E}\left[X/_{\mathfrak{G}}\right]\right|^p\leqslant \mathrm{E}\left[|X|^p/_{\mathfrak{G}}\right]$, si $1\leqslant p<\infty$. I prenent esperances, $\|\mathrm{E}\left[X/_{\mathfrak{G}}\right]\|_{L^p}\leqslant \|X\|_{L^p}$.
8) *Propietat de la Torre:* Si $\mathfrak{G}'$ és una sub-σ-àlgebra de $\mathfrak{G}$, aleshores $\mathrm{E}\left[\mathrm{E}\left[X/_{\mathfrak{G}}\right]/_{\mathfrak{G}'}\right]=\mathrm{E}\left[X/_{\mathfrak{G}'}\right]$.
9) *'Allò que és mesurable surt fora':* Si $X\in L^p$, $Y\in L^q$, on $1/p+1/q=1$, $1\leqslant p,q\leqslant\infty$, i Y és $\mathfrak{G}$-mesurable, aleshores $\mathrm{E}\left[XY/_{\mathfrak{G}}\right]=Y\,\mathrm{E}\left[X/_{\mathfrak{G}}\right]$. En particular, si X és $\mathfrak{G}$-mesurable, $\mathrm{E}\left[X/_{\mathfrak{G}}\right]=X$.
10) *'Allò que és independent desapareix':* Si $\mathfrak{G}'$ és una σ-àlgebra independent de $\mathfrak{G}$ i $\sigma\{X\}$, aleshores $\mathrm{E}\left[X/_{\mathfrak{G}\vee\mathfrak{G}'}\right]=\mathrm{E}\left[X/_{\mathfrak{G}}\right]$. En particular, si X és independent de $\mathfrak{G}$, $\mathrm{E}\left[X/_{\mathfrak{G}}\right]=\mathrm{E}[X]$.
11) *Propietat de substitució:* Siguin X_i variables aleatòries E_i-valuades, on $(E_i,\mathfrak{E}_i)$ són espais mesurables, $i=1,2$, i $\varphi\colon E_1\times E_2\to\mathbb{R}$ funció mesurable. Si X_1 és $\mathfrak{G}$-mesurable, i $h(X_1,X_2)$ és integrable, aleshores $\mathrm{E}\left[h(X_1,X_2)/_{\mathfrak{G}}\right]=\mathrm{E}\left[h(a,X_2)/_{\mathfrak{G}}\right]\big|_{a=X_1}$.

La propietat 7 ens diu, en particular, que l'operador esperança condicionada és de norma 1 de L^p en L^p, $1\leqslant p<\infty$. Això també és cert per L^∞, perquè per a mesures finites es té $\|\cdot\|_{L^p}\xrightarrow[p\to\infty]{}\|\cdot\|_{L^\infty}$.

2.4 Teorema de factorització

2.4.1 Teorema de factorització (general)

Sigui Ω un conjunt.
Sigui $(E,\mathfrak{E})$ un espai mesurable.
Siguin $X\colon\Omega\to\mathbb{R}$ i $Y\colon\Omega\to E$ dues aplicacions.
Aleshores, X és $\sigma\{Y\}$-mesurable sii existeix una funció $\varphi\colon E\to\mathbb{R}$, $(\mathfrak{E},\mathfrak{B}(\mathbb{R}))$-mesurable tal que

$$X=\varphi(Y)\quad .$$

Demostració: Posem en marxa la màquina habitual:

Suposem $X = \mathbf{1}_A$, $A \in \sigma\{Y\}$. El conjunt A es pot expressar $A = Y^{-1}(B)$, per un cert $B \in \mathfrak{E}$. Prenem $\varphi = \mathbf{1}_B$ i ja està.

Suposem $X = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$. Siguin $B_1, \ldots, B_n$ tals que $A_i = Y^{-1}(B_i)$. Prenem $\varphi = \sum_{i=1}^{n} a_i \mathbf{1}_{B_i}$.

Finalment, usem que tota funció real mesurable és límit puntual d'una successió de funcions elementals. Sigui $\{X_n\}_{n\in\mathbb{N}}$ una tal successió convergint a la funció mesurable X. Tenim que $X_n = \varphi_n(Y)$, per certes φ_n mesurables. Definim $\varphi = \lim_{n\to\infty} \varphi_n$, allà on el límit existeixi, i $\varphi = 0$ on no existeixi. Llavors,

$$X(\omega) = \lim_{n\to\infty} X_n(\omega) = \lim_{n\to\infty} \varphi_n(Y(\omega)) = \varphi(Y(\omega)) \quad . \qquad \square$$

En Williams [14], Ap. A.3.2, hi ha una demostració usant el Teorema de les Classes Monòtones.

2.4.2 Corol·lari

Si $\{Y_i\}_{i\in I}$ és una família arbitrària de funcions $\Omega \to \mathbb{R}$, aleshores $X\colon \Omega \to \mathbb{R}$ és $\sigma\{\{Y_i\}_{i\in I}\}$-mesurable sii existeix un conjunt numerable d'índexos $J \subset I$ i una funció $\varphi\colon \mathbb{R}^J \to \mathbb{R}$, $(\mathfrak{B}(\mathbb{R}^J), \mathfrak{B}(\mathbb{R}))$-mesurable, tals que

$$X = \varphi((Y_j)_{j\in J}) \quad .$$

Demostració: Considerem les composicions

$$\begin{array}{rcccl} Y_i\colon \Omega & \xrightarrow{\;Y\;} & (\mathbb{R}^I, \mathfrak{B}(\mathbb{R})^I) & \xrightarrow{\;\pi_i\;} & (\mathbb{R}, \mathfrak{B}(\mathbb{R})) \\ \omega & \longmapsto & (Y_i(\omega))_{i\in I} & \longmapsto & Y_i(\omega) \end{array}$$

En aquesta situació, donada una σ-àlgebra qualsevol en Ω, Y és mesurable sii $\forall i$, Y_i és mesurable.

A partir d'això, és immediat que $\sigma\{Y\} = \sigma\{\{Y_i\}_{i\in I}\}$.

Aplicant el Teorema de Factorització 2.4.1, X és $\sigma\{Y\}$-mesurable sii existeix $\varphi\colon \mathbb{R}^I \to \mathbb{R}$, $(\mathfrak{B}(\mathbb{R})^I, \mathfrak{B}(\mathbb{R}))$-mesurable, tal que $X = \varphi(Y)$.

Però la σ-àlgebra $\mathfrak{B}(\mathbb{R})^I$ està formada pels σ-cilindres. Per tant, $\varphi^{-1}(\mathbb{R})$ ha de ser un σ-cilindre, el que equival a dir que φ depèn d'una quantitat numerable de coordenades. O sigui, per un conjunt numerable J, $\varphi\colon \mathbb{R}^J \to \mathbb{R}$ ha de ser mesurable respecte les respectives σ-àlgebres de Borel (vegeu Observació 1.2.1). $\square$

2.4.3 Corol·lari

Si $Y\colon \Omega \to \mathbb{R}^n$, aleshores $X\colon \Omega \to \mathbb{R}$ és $\sigma\{Y\}$-mesurable sii existeix una funció mesurable $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ tal que

$$X = \varphi(Y) \quad .$$

Demostració: És conseqüència immediata del Corol·lari 2.4.2.

2.5 Esperança condicionada per variables aleatòries

2.5.1 Definició

Donada una família arbitrària de variables aleatòries $\{Y_i\}_{i\in I}$, *l'esperança d'una variable X condicionada a la família $\{Y_i\}_{i\in I}$* és l'esperança de X condicionada a la σ-àlgebra $\mathfrak{G} = \sigma\{\{Y_i\}_{i\in I}\}$. S'escriu habitualment

$$\mathrm{E}\left[{}^{X}\!/_{Y_i,\ i\in I}\right] \quad . \quad \square$$

Per exemple, una conseqüència de les propietats 10) i 11) de l'Apartat 2.3 és: Si X_1 i X_2 són independents, i h és integrable, aleshores $\mathrm{E}\left[{}^{h(X_1,X_2)}\!/_{X_1}\right] = \mathrm{E}[h(a,X_2)]_{|a=X_1}$.

En els tractats elementals de probabilitat s'intenta donar sentit a l'expressió $\mathrm{E}\left[{}^{X}\!/_{Y=y}\right]$ com un valor numèric que representa el valor mitjà de X un cop se sap que Y ha pres el valor y. Vejam com es pot fer des d'un punt de vista avançat:

Suposem que $Y\colon\Omega \to E$ és una variable aleatòria a valors en un espai mesurable $(E,\mathfrak{E})$. Com que l'esperança condicionada $\mathrm{E}\left[{}^{X}\!/_{Y}\right]$ és $\sigma\{Y\}$-mesurable, pel Teorema de Factorització 2.4.1, existeix $\varphi\colon(E,\mathfrak{E}) \to (\mathbb{R},\mathfrak{B}(\mathbb{R}))$ tal que

$$\mathrm{E}\left[{}^{X}\!/_{Y}\right] = \varphi\circ Y \quad .$$

Definim

$$\mathrm{E}\left[{}^{X}\!/_{Y=y}\right] = \varphi(y) \quad .$$

Podriem dir que $\mathrm{E}\left[{}^{X}\!/_{Y=y}\right]$ és el valor que pren la variable esperança condicionada en un ω tal que $Y(\omega)=y$. Però cal anar amb compte perquè, com es veu en la demostració del Teorema 2.4.1, la funció φ és única llevat de conjunts de probabilitat P_Y nul·la. Per tant, de fet els valors concrets $\varphi(y)$ no estan en general ben definits. Només tenen un sentit global. Globalment, la funció φ queda caracteritzada per la igualtat

$$\begin{aligned}\int_{Y^{-1}(B)} X(\omega)\,P(d\omega) &= \int_{Y^{-1}(B)} \mathrm{E}\left[{}^{X}\!/_{Y}\right](\omega)\,P(d\omega)\\ &= \int_{Y^{-1}(B)} \big(\varphi\circ Y\big)(\omega)\,P(d\omega) = \int_B \varphi(y)\,P_Y(dy) \quad ,\end{aligned}$$

pel Teorema de la Mesura Imatge.

Resumint, són vàlides les expressions:

$$\begin{aligned}&1) \qquad \int_A X\,dP = \int_A \mathrm{E}\left[{}^{X}\!/_{Y}\right] dP \quad , \quad \forall A\in\sigma\{Y\} \quad .\\ &2) \qquad \int_{Y^{-1}(B)} X\,dP = \int_B \mathrm{E}\left[{}^{X}\!/_{Y=y}\right] P_Y(dy) \quad , \quad \forall B\in\mathfrak{B}(\mathbb{R}) \quad .\end{aligned}$$

2.6 Probabilitat condicionada

2.6.1 Definició

Sigui $(\Omega, \mathfrak{F}, P)$ un espai de probabilitat.
Sigui $\mathfrak{G}$ una sub-σ-àlgebra de $\mathfrak{F}$.
Sigui $A \in \mathfrak{F}$.
Anomenem *probabilitat de A condicionada a $\mathfrak{G}$* a la variable aleatòria

$$P\big({}^{A}/_{\mathfrak{G}}\big) := \mathrm{E}\left[\mathbf{1}_{A}/_{\mathfrak{G}}\right] \quad . \quad \Box$$

Pensem ara en l'esperança condicionada com una funció de conjunt que està determinada excepte conjunts de probabilitat zero. És fàcil veure que es compleixen les propietats següents:

1) $P\big({}^{\emptyset}/_{\mathfrak{G}}\big) = 0$, q.s.
2) $P\big({}^{\Omega}/_{\mathfrak{G}}\big) = 1$, q.s.
3) $A \subset B \Rightarrow P\big({}^{A}/_{\mathfrak{G}}\big) \leqslant P\big({}^{B}/_{\mathfrak{G}}\big)$, q.s.
4) $\forall \{A_n\}_{n\in\mathbb{N}} \subset \mathfrak{G}$, disjunts dos a dos, $P\big(\bigcup_{n=1}^{\infty} A_n/_{\mathfrak{G}}\big) = \sum_{n=1}^{\infty} P\big({}^{A_n}/_{\mathfrak{G}}\big)$, q.s.

Podem pensar que tenim l'aplicació

$$\begin{aligned} p\colon \Omega \times \mathfrak{F} &\longrightarrow \mathbb{R} \\ (\omega, A) &\longmapsto p(\omega, A) = P\big({}^{A}/_{\mathfrak{G}}\big)(\omega) \end{aligned}$$

i voldriem que $p(\omega, \cdot)$ fos una probabilitat en $(\Omega, \mathfrak{F})$, si més no P-q.p.t. ω. Ho és? La pregunta precisa és:

És possible escollir, per a cada $A \in \mathfrak{F}$, un representant de la variable aleatòria $P\big({}^{A}/_{\mathfrak{G}}\big)$ de manera que $p(\omega, \cdot)$ sigui una probabilitat en $(\Omega, \mathfrak{F})$?.

També es pot formular la pregunta d'aquesta altra manera: És possible escollir, per a cada $\omega \in \Omega$, una probabilitat $p(\omega, \cdot)$ tal que es tingui la igualtat

$$\mathrm{E}\left[{}^{X}/_{\mathfrak{G}}\right](\omega) = \int_{\Omega} X(\omega')\, p(\omega, d\omega') \quad ?$$

En general, no; gairebé sempre, sí.

La dificultat és que en la propietat 4) anterior la igualtat pot fallar en un conjunt de probabilitat zero associat a la família $\{A_n\}_{n\in\mathbb{N}}$. Però de famílies numerables n'hi ha una infinitud no numerable, en general. Per tant, no podem assegurar d'entrada que, llevat d'un conjunt de probabilitat zero, se satisfagui la igualtat per a totes les possibles famílies.

Vegeu en Williams [13], pàg. 56, un exemple de situació en la qual no és possible fer que $p(\omega, \cdot)$ sigui una probabilitat P-q.p.t. ω. (No és fàcil.) Un resultat positiu general és el del Teorema 2.6.3.

2.6.2 Definició

Sigui $(\Omega, \mathfrak{F}, P)$ un espai de probabilitat.
Sigui $\mathfrak{G}$ una sub-σ-àlgebra de $\mathfrak{F}$.

Una *probabilitat condicionada de P a $\mathfrak{G}$* és una aplicació

$$\begin{aligned} p\colon \Omega \times \mathfrak{F} &\longrightarrow \mathbb{R} \\ (\omega, A) &\longmapsto p(\omega, A) \end{aligned}$$

tal que

1) $\forall A \in \mathfrak{F}$,

$$\begin{aligned} p(\cdot, A)\colon \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto p(\omega, A) \end{aligned}$$

   és una versió de la probabilitat de A condicionada a $\mathfrak{G}$, és a dir:

   a) $p(\cdot, A)$ és $\mathfrak{G}$-mesurable.

   b) $\forall G \in \mathfrak{G}$, $P(A \cap G) = \displaystyle\int_G \mathbf{1}_A \, dP = \int_G p(\cdot, A)(\omega)\, P(d\omega)$.

2) P-q.p.t. $\omega \in \Omega$,

$$\begin{aligned} p(\omega, \cdot)\colon \mathfrak{F} &\longrightarrow \mathbb{R} \\ A &\longmapsto p(\omega, A) \end{aligned}$$

   és una probabilitat en $(\Omega, \mathfrak{F})$.

2.6.3 Teorema

Sigui Ω un espai polonès (i.e. mètric, separable i complet).
Siguin $\mathfrak{F} = \mathfrak{B}(\Omega)$, i P una probabilitat en $(\Omega, \mathfrak{F})$.
Sigui $\mathfrak{G}$ una sub-σ-àlgebra de $\mathfrak{F}$.
Aleshores:

1) *Existeix una probabilitat condicionada de P a $\mathfrak{G}$.*
2) *Si p i $\tilde{p}$ són dues d'elles, $p(\omega, \cdot) = \tilde{p}(\omega, \cdot)$ excepte potser en un conjunt de $\mathfrak{G}$ de probabilitat zero.*

Podeu veure una demostració d'aquest Teorema en Stroock–Varadhan [11] (Teorema 1.1.6).

Encara podem considerar una situació una mica més general:

2.6.4 Teorema

Sigui $(\Omega, \mathfrak{F}, P)$ un espai de probabilitat, i sigui $Y\colon \Omega \to \Omega'$ una funció mesurable de $(\Omega, \mathfrak{F})$ en un espai mesurable $(\Omega', \mathfrak{F}')$, on Ω' és polonès i $\mathfrak{F}' = \mathfrak{B}(\Omega')$.
Sigui $\mathfrak{G}$ una sub-σ-àlgebra de $\mathfrak{F}$.
Aleshores, existeix una "llei condicionada" de Y a $\mathfrak{G}$. Amb precisió:
Existeix una aplicació

$$\begin{aligned} p\colon \Omega \times \mathfrak{F}' &\longrightarrow \mathbb{R} \\ (\omega, B) &\longmapsto p(\omega, B) \end{aligned}$$

tal que

1) *$\forall B \in \mathfrak{F}'$,*

$$\begin{aligned} p(\cdot, B)\colon \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto p(\omega, B) \end{aligned}$$

   satisfà

   a) *$p(\cdot, B)$ és $\mathfrak{G}$-mesurable.*

   b) *$\forall G \in \mathfrak{G}$, $P(\{Y \in B\} \cap G) = \displaystyle\int_G \mathbf{1}_{\{Y \in B\}} \, dP = \int_G p(\cdot, \{Y \in B\})(\omega)\, P(d\omega)$.*

2) P-q.p.t. $\omega \in \Omega$,

$$\begin{aligned} p(\omega,\cdot)\colon \mathfrak{F}' &\longrightarrow \mathbb{R} \\ B &\longmapsto p(\omega,B) \end{aligned}$$

és una probabilitat en $(\Omega',\mathfrak{F}')$. □

Demostració: Vegeu per exemple Ash [2], Teoremes 6.6.4, 6.6.5, i 6.6.6. □

En particular, del Teorema 2.6.4 es dedueix el Teorema 2.6.3, si prenem $(\Omega,\mathfrak{F}) = (\Omega',\mathfrak{F}')$ i Y igual a la identitat. Naturalment, $\mathbb{R}$ i $\mathbb{R}^n$ són espais polonesos, i el teorema es pot aplicar en aquest cas (lleis condicionades de variables i vectors aleatoris).

L'esperança condicionada es pot calcular a través d'una probabilitat condicionada:

2.6.5 Proposició

Sigui $X \in L^1(\Omega,\mathfrak{F},P)$.
Sigui p una probabilitat condicionada de P a la σ-àlgebra $\mathfrak{G} \subset \mathfrak{F}$.
Aleshores, P-q.p.t. $\omega \in \Omega$, X és integrable respecte $p(\omega,\cdot)$ i

$$\int_\Omega X(\omega')\, p(\omega,d\omega') = \mathrm{E}\left[{}^{X}\!/_{\mathfrak{G}}\right](\omega) \quad .$$

Demostració: Es fa usant la maquinària: Indicadors, elementals, positives, arbitràries. □

Introduint una petita hipòtesi més a les situacions generals considerades, treurem encara més conseqüències sobre la probabilitat condicionada.

Es diu que una σ-àlgebra $\mathfrak{G}$ és *numerablement generada* (o també *separable*) si existeix una família numerable $\{A_n\}_{n\in\mathbb{N}} \subset \mathfrak{G}$ tal que $\sigma\{\{A_n\}_{n\in\mathbb{N}}\} = \mathfrak{G}$.

Observació prèvia al teorema següent: Si $\mathfrak{F}$ és la σ-àlgebra de Borel d'un espai polonès, es pot demostrar que és numerablement generada. Però una σ-àlgebra numerablement generada pot contenir sub-σ-àlgebres que no ho siguin.

2.6.6 Teorema

En la situació del Teorema 2.6.3, suposem a més que $\mathfrak{G}$ és numerablement generada.
Aleshores, es pot escollir l'aplicació $\omega \mapsto p(\omega,\cdot)$ de manera que

$$p(\omega,A) = \mathbf{1}_A(\omega) \quad , \quad \forall \omega \in \Omega, \quad \forall A \in \mathfrak{G} \quad . \tag{2.6.1}$$

Demostració: La idea és: Sabem que, si $A \in \mathfrak{G}$,

$$p(\omega,A) = P\left({}^{A}\!/_{\mathfrak{G}}\right)(\omega) = \mathbf{1}_A(\omega) \quad , \quad q.s.$$

Tindrem aquesta igualtat per a tot A pertanyent a una àlgebra de cardinalitat numerable que genera $\mathfrak{G}$. La unió dels conjunts de mesura zero on falla la igualtat torna a ser de mesura zero, i per tant,

$$p(\omega,A) = \mathbf{1}_A(\omega) \quad ,$$

per a tot A d'aquesta àlgebra, excepte en un conjunt de ω's de $\mathfrak{G}$ i probabilitat zero.

Finalment, la igualtat s'estén a la σ-àlgebra generada $\mathfrak{G}$, i el q.s. es pot treure escollint una versió pertinent. $\square$

En particular, si $A(\omega)$ és *l'àtom* contenint ω, que es defineix com el conjunt

$$A(\omega) := \cap\{A \in \mathfrak{G} : \ \omega \in A\} \quad ,$$

i que es pot demostrar que pertany a $\mathfrak{G}$ si aquesta σ-àlgebra és numerablement generada, es té que

$$p(\omega, A(\omega)) = 1 \quad .$$

2.6.7 Definició

Si p és una probabilitat condicionada que compleix (2.6.1), és diu que p és una *versió regular de la probabilitat condicionada de P a $\mathfrak{G}$*. $\square$

Conseqüència de tot això és que sempre existeix una versió regular de la "llei" d'un vector aleatori condicionada a un altre vector aleatori.

3. Motivació dels processos de Markov

3.1 Motivació

Hem vist que la llei d'un procés estocàstic queda determinada per la família de lleis en dimensió finita. Considerem els dos casos particulars següents:

Totes les variables aleatòries del procés són independents entre si. Això equival a dir, per la definició d'independència de variables aleatòries, que les lleis en dimensió finita del procés són probabilitats producte en $\mathbb{R}^n$, $p_1 \times \cdots \times p_n$.

Un altre cas molt particular és aquell en que les variables estan relacionades per una equació diferencial ordinària (i la primera variable X_0 determinaria totes les altres variables), o més generalment, per qualsevol relació funcional del tipus $X_t = F(t, X_0)$.

Això són dos extrems. Els casos interessants solen estar entre mig. Al primer extrem, l'atzar intervé tan fortament a cada instant de temps que el passat no importa en absolut per saber que està passant ara o què passarà en el futur (intuïtivament, les trajectòries del procés seran molt discontínues). Al segon extrem, l'atzar només intervé al principi (o fins i tot ni això). Després, tot el futur està determinat totalment pel valor de la variable inicial.

Els cassos interessants són aquells en què hi ha una intervenció de l'atzar en cada instant de temps, però també hi ha alguna influència del passat (per exemple, imagineu un procés amb trajectòries contínues).

Què vol dir influència del passat? En principi, cada variable X_t d'un procés té una llei. Però llevat de les situacions d'indepèndencia, el coneixement dels valors que han pres altres variables en instants anteriors (o posteriors) a t modifica el que podem esperar de la variable X_t. Aquesta modificació és el que formalitzem amb les probabilitats condicionades. Suposem per exemple que coneixem els valors de les variables $X_{t_1}, \dots, X_{t_{j-1}}$. Aleshores, el que interessa de X_{t_j} és la seva llei condicionada a aquests valors, és a dir, les quantitats

$$P\big(X_{t_j} \in B \big/ X_{t_1} = x_1, \dots, X_{t_{j-1}} = x_{j-1}\big) \quad , \quad B \in \mathfrak{B}(\mathbb{R}) \quad .$$

La idea de procés de Markov és la de que tota la informació de què es disposa sobre el passat es pot resumir en la informació que conté la última de les variables del passat que considerem:

$$P\big(X_{t_j} \in B \big/ X_{t_1} = x_1, \dots, X_{t_{j-1}} = x_{j-1}\big) = P\big(X_{t_j} \in B \big/ X_{t_{j-1}} = x_{j-1}\big) \quad .$$

Ve a ser l'anàleg aleatori del que passa amb sistemes deterministes governats per una equació diferencial ordinària amb unicitat respecte el problema de Cauchy: Un valor donat en un instant de temps determina el sistema en instant posteriors; no es necessita informació sobre on estava abans el sistema. De fet, també el passat queda determinat, i aquí passa el mateix:

$$P\big(X_{t_j}\in B / X_{t_{j+1}}=x_{j+1},\dots,X_{t_n}=x_n\big)=P\big(X_{t_j}\in B / X_{t_{j+1}}=x_{j+1}\big) \quad .$$

(Ho veurem més endavant, un cop ben establertes les definicions.) No importa com formulem de manera rigorosa el concepte de procés de Markov (que es pot fer de maneres diverses equivalents), aquesta és la idea que cal tenir sempre en ment.

3.2 Exemples: Processos de Wiener, d'Ornstein–Uhlenbeck i de Poisson

Hi ha moltes maneres més o menys equivalents de definir el procés de Wiener. Per fixar idees, utilitzarem la següent:

3.2.1 Definició

Un procés estocàstic real $W\equiv\{W_t,\ t\in\mathbb{R}^+\}$ és un *procés de Wiener* si

1) És continu.
2) Té increments independents del passat.
3) Per a tots $0\leqslant s<t$, W_t-W_s és una variable aleatòria gaussiana amb esperança zero i variància $\sigma^2(t-s)$, per algun $\sigma^2>0$.

Un procés de Wiener *comença en* $x\in\mathbb{R}$ si $W_0=x$, q.s.
Si a més, $\sigma^2=1$, W s'anomena *procés de Wiener estàndar començant en* $x\in\mathbb{R}$. □

La funció

$$f(x;t,y)=(2\pi t)^{-1/2}\exp\{-(y-x)^2/2t\} \quad , \quad t>0 \tag{3.2.1}$$

està relacionada amb el procés de Wiener. És la seva *funció de densitat de transició*. Això s'interpreta en el sentit següent: Si el procés està a l'estat x a l'instant s, aleshores la probabilitat que el procés es trobi en un petit interval de longitud dy al voltant de y en un instant posterior r és

$$f(x;r-s,y)\,dy \quad .$$

La independència del passat i el futur del procés quan es coneix el seu estat en el present és la *propietat de Markov* del procés de Wiener.

3.2.2 Exercici

Usant la forma explícita de p (3.2.1) es comprova que

$$\int_{\mathbb{R}} f(x;s,y)p(y;t,z)\,dy=p(x;s+t,z) \quad , \tag{3.2.2}$$

$x,z\in\mathbb{R}$ i $s,t>0$. És un cas particular de la fórmula (5.1.2), més endavant. □

La igualtat (3.2.2) es pot expressar de la manera següent: Definim

$$T_t(g):=\int_{\mathbb{R}} f(x;t,y)g(y)\,dy \quad ,$$

aplicació que tranforma funcions en funcions (oblidem de moment en quin espai de funcions l'hem de definir). Llavors (3.2.2) implica

$$T_s \circ T_t = T_{s+t}$$

Es diu que la família d'operadors $\{T_t\}_{t\in\mathbb{R}^+}$ és un *semigrup*. Veurem més endavant que això és un fet general dels processos de Markov, que ens permetrà usar la teoria de semigrups per estudiar-los.

3.2.3 Definició

Un procés estocàstic real $X \equiv \{X_t,\ t \in \mathbb{R}^+\}$ és un *procés de Ornstein–Uhlenbeck* amb paràmetre $\gamma > 0$ si

1) $X_0 \sim N(0,1)$.
2) $\mathrm{E}[X_t] = 0$, $\mathrm{E}[X_s X_t] = e^{-\gamma(t-s)}$, $\forall s,t \in \mathbb{R}^+$. □

El procés d'Ornstein–Uhlenbeck té funció de densitat de transició

$$f(x;t,y) = \frac{1}{\sqrt{2\pi(1-e^{-2\gamma t})}} \exp\Big\{ -\frac{(y-xe^{-\gamma t})^2}{2(1-e^{-2\gamma t})}\Big\} \quad ,$$

amb la mateixa interpretació que per al procés de Wiener.

3.2.4 Definició

Un procés estocàstic real $X \equiv \{X_t,\ t \in \mathbb{R}^+\}$ és un *procés de Poisson* amb intensitat $\lambda > 0$ si

1) $X_0 = 0$.
2) $X_t - X_s \sim \mathrm{Pois}(\lambda(t-s))$. $\forall s,t \in \mathbb{R}^+$.
3) Té increments independents del passat. □

En aquest cas no hi ha densitat de transició, perquè el procés pren valors en $\mathbb{N}$, però és un procés de Markov i es podran definir unes “probabilitats de transició”, encara que no tinguin densitat respecte la mesura de Lebesgue.

3.3 Més idees intuïtives

Suposem per simplificar que les lleis en dimensió finita d'un procés $\{X_t,\ t \in \mathbb{R}^+\}$ són absolutament contínues respecte la mesura de Lebesgue. Notem $f_{X_{t_1},\dots,X_{t_n}}$ la densitat de $(X_{t_1},\dots,X_{t_n})$ i $f_{X_{t_1},\dots,X_{t_n}|_{X_{s_1}=y_1,\dots,X_{s_m}=y_m}}$ la densitat condicionada de $(X_{t_1},\dots,X_{t_n})$ donats $X_{s_1}=y_1,\dots,X_{s_m}=y_m$.

La propietat de Markov ens diu que si $t_1 < \cdots < t_n$,

$$f_{X_{t_n}|_{X_{t_1}=y_1,\dots,X_{t_{n-1}}=y_{n-1}}}(x) = f_{X_{t_n}|_{X_{t_{n-1}}=y_{n-1}}}(x) \quad .$$

L'ús d'aquesta propietat simplifica coses. Per exemple, la fórmula general

$$f_{X_{t_1},X_{t_2},X_{t_3}}(y_1,y_2,y_3) = f_{X_{t_3}|_{X_{t_1}=y_1,X_{t_2}=y_2}}(y_3)\cdot f_{X_{t_2}|_{X_{t_1}=y_1}}(y_2)\cdot f_{X_{t_1}}(y_1)$$

s'expressarà, per un procés de Markov,

$$f_{X_{t_1},X_{t_2},X_{t_3}}(y_1,y_2,y_3) = f_{X_{t_3}|_{X_{t_2}=y_2}}(y_3)\cdot f_{X_{t_2}|_{X_{t_1}=y_1}}(y_2)\cdot f_{X_{t_1}}(y_1) \quad ,$$

i en general

$$f_{X_{t_1},\dots,X_{t_n}}(y_1,\dots,y_n) =$$
$$f_{X_{t_n}|_{X_{t_{n-1}}=y_{n-1}}}(y_n)\cdot f_{X_{t_{n-1}}|_{X_{t_{n-2}}=y_{n-2}}}(y_{n-1})\cdot\ldots\cdot f_{X_{t_2}|_{X_{t_1}=y_1}}(y_2)\cdot f_{X_{t_1}}(y_1) \quad .$$

Intuïtivament, veiem que la densitat de la primera variable del procés i les densitats condicionades $f_{X_t|_{X_s=\cdot}}$ determinen la llei del procés. Aquest és un fet general que veurem en el Capítol 5.

4. Processos de Markov

4.1 Independència condicionada

Recordem que els espais $L^p(\Omega, \mathfrak{F}, P)$, $1 \leqslant p < \infty$ (abreviat $L^p(\Omega)$ o L^p quan no hi ha possibilitat de confusió) estan formats per classes de variables aleatòries X mòdul la relació d'equivalència P-q.s. que compleixen $\int_\Omega |X|^p \, dP < \infty$. Són espais de Banach amb la norma $\|X\|_{L^p} := \Big(\int_\Omega |X|^p \, dP \Big)^{1/p}$. L'espai L^2 és de Hilbert amb el producte escalar $\langle X, Y \rangle := \int_\Omega XY \, dP$.

També es poden considerar de la mateixa manera els espais L^p amb $0 \leqslant p < 1$, que són espais mètrics complets amb la distància definida per $d(X,Y) := \int_\Omega |X - Y|^p \, dP$ (si $p > 0$) i per la distància definida per la convergència en probabilitat si $p = 0$ (L^0 és l'espai vectorial de totes les variables aleatòries), però en aquests casos la relació entre la estructura algebraica i l'estructura topològica és molt dolenta.

Es defineix L^∞ com l'espai de classes d'equivalència de variables aleatòries tals que $\|X\|_{L^\infty} := \operatorname{ess\,sup} |X| < \infty$. També és de Banach. Si $p \leqslant q$, aleshores $L^q \subset L^p$. Es té:

$$L^0 \supset \cdots \supset L^p \supset L^q \supset \cdots \supset L^\infty \cdots$$

4.1.1 Proposició

Sigui $(\Omega, \mathfrak{F}, P)$ un espai de probabilitat.
Siguin $\mathfrak{A}_1$, $\mathfrak{A}_2$, $\mathfrak{B}$ sub-σ-àlgebres de $\mathfrak{F}$.
Aleshores són equivalents:

i) $\forall A_1 \in \mathfrak{A}_1$, $\forall A_2 \in \mathfrak{A}_2$,

$$P\big(A_1 \cap A_2 / \mathfrak{B}\big) = P\big(A_1/\mathfrak{B}\big) P\big(A_2/\mathfrak{B}\big) \quad , \quad q.s.$$

ii) $\forall \xi_1 \in L^2(\Omega, \mathfrak{A}_1, P)$, $\forall \xi_2 \in L^2(\Omega, \mathfrak{A}_2, P)$,

$$\mathrm{E}\big[\xi_1\xi_2/\mathfrak{B}\big] = \mathrm{E}\big[\xi_1/\mathfrak{B}\big]\, \mathrm{E}\big[\xi_2/\mathfrak{B}\big] \quad , \quad q.s.$$

iii) $\forall \xi_1, \xi_2$ *pertanyents a un sistema total (Total=generen un subspai vectorial dens) de* $L^2(\Omega, \mathfrak{A}_1, P)$ *i* $L^2(\Omega, \mathfrak{A}_2, P)$*, respectivament,*

$$\mathrm{E}\left[\xi_1\xi_2/_{\mathfrak{B}}\right] = \mathrm{E}\left[\xi_1/_{\mathfrak{B}}\right]\mathrm{E}\left[\xi_2/_{\mathfrak{B}}\right] \quad , \quad q.s.$$

4.1.2 Exercici

La proposició anterior és certa.

4.1.3 Definició

En les situacions de la Proposició 4.1.1, diem que $\mathfrak{A}_1$ i $\mathfrak{A}_2$ són condicionalment independents donada $\mathfrak{B}$ i ho notem

$$\mathfrak{A}_1 \underset{\mathfrak{B}}{\perp\!\!\!\perp} \mathfrak{A}_2 \quad . \quad \square$$

Es nota habitualment per $\mathfrak{A} \vee \mathfrak{B}$ la σ-àlgebra generada per $\mathfrak{A} \cup \mathfrak{B}$, que està generada per la família de conjunts $\{A \cap B : \ A \in \mathfrak{A}, \ B \in \mathfrak{B}\}$.

4.1.4 Proposició

Si $\mathfrak{A}_1'$ *i* $\mathfrak{A}_2'$ *són sub-*σ*-àlgebres de* $\mathfrak{A}_1 \vee \mathfrak{B}$ *i* $\mathfrak{A}_2 \vee \mathfrak{B}$ *respectivament, aleshores*

$$\mathfrak{A}_1 \underset{\mathfrak{B}}{\perp\!\!\!\perp} \mathfrak{A}_2 \Rightarrow \mathfrak{A}_1' \underset{\mathfrak{B}}{\perp\!\!\!\perp} \mathfrak{A}_2' \quad .$$

Demostració: Notem $\tilde{\mathfrak{A}}_1 = \mathfrak{A}_1 \vee \mathfrak{B}$. Un sistema total de $L^2(\Omega, \tilde{\mathfrak{A}}_1, P)$ és el format per les variables de la forma $\tilde{\xi}_1 = \xi_1\eta_1$, amb $\xi_1 = \mathbf{1}_A$ i $\eta_1 = \mathbf{1}_B$, $A \in \mathfrak{A}_1$, $B \in \mathfrak{B}$, i anàlogament per $\tilde{\mathfrak{A}}_2 = \mathfrak{A}_2 \vee \mathfrak{B}$ i $L^2(\Omega, \tilde{\mathfrak{A}}_2, P)$.

Llavors,

$$\begin{aligned} \mathrm{E}\left[\tilde{\xi}_1\tilde{\xi}_2/_{\mathfrak{B}}\right] &= \eta_1\eta_2\,\mathrm{E}\left[\xi_1\xi_2/_{\mathfrak{B}}\right] \\ &= \eta_1\eta_2\,\mathrm{E}\left[\xi_1/_{\mathfrak{B}}\right]\mathrm{E}\left[\xi_2/_{\mathfrak{B}}\right] = \mathrm{E}\left[\tilde{\xi}_1/_{\mathfrak{B}}\right]\mathrm{E}\left[\tilde{\xi}_2/_{\mathfrak{B}}\right] \quad , \end{aligned}$$

i obtenim $\tilde{\mathfrak{A}}_1 \underset{\mathfrak{B}}{\perp\!\!\!\perp} \tilde{\mathfrak{A}}_2$.

Que podem prendre sub-σ-àlgebres de $\tilde{\mathfrak{A}}_1$ i $\tilde{\mathfrak{A}}_2$ i es manté l'anterior independència condicionada és obvi.

4.1.5 Definició

Una seqüència de σ-àlgebres $\mathfrak{A}_1$, $\mathfrak{B}$, $\mathfrak{A}_2$ és *markoviana* si i només si

$$P\left(A_2/_{\mathfrak{A}_1 \vee \mathfrak{B}}\right) = P\left(A_2/_{\mathfrak{B}}\right) \quad , \quad \forall A_2 \in \mathfrak{A}_2 \quad . \quad \square \tag{4.1.1}$$

Anàlogament a les equivalències de la Proposició 4.1.1, es demostra que (4.1.1) és equivalent a

$$\forall \xi_2 \in L^2(\Omega, \mathfrak{A}_2, P), \quad \mathrm{E}\left[\xi_2/_{\mathfrak{A} \vee \mathfrak{B}}\right] = \mathrm{E}\left[\xi_2/_{\mathfrak{B}}\right] \quad ,$$

i el mateix substituint '$\forall \xi_2$' per '$\forall \xi_2$ pertanyent a un sistema total de $L^2(\Omega, \mathfrak{A}_2, P)$'.

4.1.6 Proposició

$\mathfrak{A}_1 \underset{\mathfrak{B}}{\perp\!\!\!\perp} \mathfrak{A}_2 \Leftrightarrow$ *la seqüència* $\mathfrak{A}_1$*,* $\mathfrak{B}$*,* $\mathfrak{A}_2$ *és markoviana.*

Demostració:
1) Cas particular $\mathfrak{B} \subset \mathfrak{A}_1$:
$\Leftarrow$) Si $\xi_1 \in L^2(\Omega, \mathfrak{A}_1, P)$ i $\xi_2 \in L^2(\Omega, \mathfrak{A}_2, P)$, tindrem

$$\begin{aligned} \mathrm{E}\left[\xi_1\xi_2/\mathfrak{B}\right] &= \mathrm{E}\left[\xi_1\xi_2/\mathfrak{A}_1\right] \\ &= \mathrm{E}\left[\xi_1\,\mathrm{E}\left[\xi_2/\mathfrak{A}_1\right]/\mathfrak{B}\right] \quad . \end{aligned} \tag{4.1.2}$$

Però, per hipòtesi,

$$\mathrm{E}\left[\xi_2/\mathfrak{A}_1\right] = \mathrm{E}\left[\xi_2/\mathfrak{B}\right] \quad .$$

Per tant, (4.1.2) és igual a

$$\mathrm{E}\left[\xi_1\,\mathrm{E}\left[\xi_2/\mathfrak{B}\right]/\mathfrak{B}\right] = \mathrm{E}\left[\xi_1/\mathfrak{B}\right]\mathrm{E}\left[\xi_2/\mathfrak{B}\right] \quad .$$

$\Rightarrow$) De la independència condicionada es dedueix que, per $\xi_1 \in L^2(\Omega, \mathcal{A}_1, P)$ i $\xi_2 \in L^2(\Omega, \mathcal{A}_2, P)$,

$$\mathrm{E}\left[\xi_1\xi_2 - \xi_1\,\mathrm{E}\left[\xi_2/\mathfrak{B}\right]/\mathfrak{B}\right] = 0 \quad ,$$

d'on resulta que les variables ξ_1 i $\xi_2 - \mathrm{E}\left[\xi_2/\mathfrak{B}\right]$ són ortogonals a l'espai de Hilbert $L^2(\Omega, \mathfrak{F}, P)$. Com que $\mathfrak{B} \subset \mathfrak{A}_1$, la variable $\mathrm{E}\left[\xi_2/\mathfrak{B}\right]$ és la projecció de ξ_2 sobre $L^2(\Omega, \mathfrak{A}_1, P)$. Per tant,

$$\mathrm{E}\left[\xi_2/\mathfrak{B}\right] = \mathrm{E}\left[\xi_2/\mathfrak{A}_1\right]$$

i la seqüència és markoviana.

2) Cas general:
Clarament,

$$\mathfrak{A}_1 \underset{\mathfrak{B}}{\perp\!\!\!\perp} \mathfrak{A}_2 \Leftrightarrow \mathfrak{A}_1 \vee \mathfrak{B} \underset{\mathfrak{B}}{\perp\!\!\!\perp} \mathfrak{A}_2$$

per la Proposició 4.1.4. Pel cas anterior, això és equivalent també a què $\mathfrak{A}_1 \vee \mathfrak{B}$, $\mathfrak{B}$, $\mathfrak{A}_2$ és markoviana, és a dir, a què

$$P\left(A_2/\mathfrak{A}_1 \vee \mathfrak{B} \vee \mathfrak{B}\right) = P\left(A_2/\mathfrak{B}\right) \quad , \quad \forall A_2 \in \mathfrak{A}_2 \quad ,$$

que equival al seu torn a dir que $\mathfrak{A}_1$, $\mathfrak{B}$, $\mathfrak{A}_2$ és markoviana. □

La independència condicionada és un concepte delicat pel que fa a la manipulació de la σ-àlgebra que condiciona. No és cert que $\mathfrak{A}_1 \underset{\mathfrak{B}}{\perp\!\!\!\perp} \mathfrak{A}_2$ impliqui la mateixa propietat substituint $\mathfrak{B}$ per qualsevol σ-àlgebra més petita o més gran que $\mathfrak{B}$. Per exemple, $\mathfrak{A}_1 \underset{\mathfrak{A}_1}{\perp\!\!\!\perp} \mathfrak{A}_2$ és cert sempre, mentre que $\mathfrak{A}_1 \underset{\{\emptyset,\Omega\}}{\perp\!\!\!\perp} \mathfrak{A}_2$, equivalent a dir que $\mathfrak{A}_1$ i $\mathfrak{A}_2$ són independents, no és cert en general. Per altra banda, per exemple, si X_1 i X_2 són variables independents, no és cert en general que $\sigma\{X_1\} \underset{\sigma\{X_1+X_2\}}{\perp\!\!\!\perp} \sigma\{X_2\}$.

4.2 Processos de Markov

4.2.1 Definició

Sigui $(\Omega, \mathfrak{F}, P)$ un espai de probabilitat.
Sigui $\{X_t,\ t \in \mathbb{R}^+\}$ un procés estocàstic real sobre $(\Omega, \mathfrak{F}, P)$.

Diem que $\{X_t,\ t \in \mathbb{R}^+\}$ és un *procés de Markov* sii

$$\forall t \in \mathbb{R}^+, \sigma\{X_s,\ s \leqslant t\} \underset{\sigma\{X_t\}}{\perp\!\!\!\perp} \sigma\{X_s,\ s \geqslant t\} \quad . \qquad \square$$

Equivalentment, la seqüència $\sigma\{X_s,\ s \leqslant t\}$, $\sigma\{X_t\}$, $\sigma\{X_s,\ s \geqslant t\}$ és markoviana. $\square$

Per definició, la independència de variables aleatòries és la independència de les σ-àlgebres que generen. Anàlogament amb la independència condicionada. La condició de la definició anterior es pot escriure doncs

$$\forall t \in \mathbb{R}^+, \{X_s,\ s \leqslant t\} \underset{X_t}{\perp\!\!\!\perp} \{X_s,\ s \geqslant t\} \quad .$$

(és purament notació).

El fet que la seqüència $\sigma\{X_s,\ s \leqslant t\}$, $\sigma\{X_t\}$, $\sigma\{X_s,\ s \leqslant t\}$ sigui markoviana reflecteix clarament la idea que, si estem a l'instant t, la informació que dóna el passat ($\{X_s,\ s \leqslant t\}$) sobre el futur ($\{X_s,\ s \geqslant t\}$) està continguda en el present (X_t). La definició en termes d'independència condicionada ens porta a formular aquesta mateixa idea dient que el passat i el futur són independents si coneixem el present.

És clar de la Proposició 4.1.6 que si $\mathfrak{A}_1$, $\mathfrak{B}$, $\mathfrak{A}_2$ és markoviana, aleshores $\mathfrak{A}_2$, $\mathfrak{B}$, $\mathfrak{A}_1$ també és markoviana. El paper del passat i el futur és per tant, en un procés de Markov, simètric. També és pot dir que la informació que dóna el futur sobre el passat està continguda en el present.

És clar que el que tenim fins ara implica que: $\forall \psi : \mathbb{R} \to \mathbb{R}$ mesurable i acotada, $\forall t \in \mathbb{R}^+$, $\forall h > 0$,

$$\mathrm{E}\left[\psi(X_{t+h}) / \{X_s,\ s \leqslant t\}\right] = \mathrm{E}\left[\psi(X_{t+h}) / X_t\right] \quad ,$$

i d'aquí que $\forall B \in \mathfrak{B}(\mathbb{R})$, $\forall t \in \mathbb{R}^+$, $\forall h > 0$,

$$P\big(X_{t+h} \in B / \{X_s,\ s \leqslant t\}\big) = P\big(X_{t+h} \in B / X_t\big) \quad .$$

La qüestió interessant és que resulta que el recíproc és cert, i això redueix realment la feina de comprovar la propietat de Markov.

Una segona reducció apreciable és la següent: De les propietats de l'esperança condicionada, si X és Markov, per $s_1 < \cdots < s_n < t$, es té

$$\begin{aligned}
&\mathrm{E}\left[\psi(X_{t+h}) / X_{s_1}, \ldots, X_{s_n}, X_t\right] \\
&= \mathrm{E}\Big[\mathrm{E}\left[\psi(X_{t+h}) / \{X_s,\ s \leqslant t\}\right] / X_{s_1}, \ldots, X_{s_n}, X_t\Big] \\
&= \mathrm{E}\Big[\mathrm{E}\left[\psi(X_{t+h}) / X_t\right] / X_{s_1}, \ldots, X_{s_n}, X_t\Big] \\
&= \mathrm{E}\left[\psi(X_{t+h}) / X_t\right] \quad ,
\end{aligned}$$

$\forall \psi : \mathbb{R} \to \mathbb{R}$ mesurable i acotada, $\forall t \in \mathbb{R}^+$, $\forall h > 0$. En particular, $\forall t \in \mathbb{R}^+$, $\forall h > 0$, $\forall B \in \mathfrak{B}(\mathbb{R})$,

$$P\big(X_{t+h} \in B / X_{s_1} \ldots, X_{s_n}, X_t\big) = P\big(X_{t+h} \in B / X_t\big) \quad .$$

Com abans, el recíproc també és cert, i arribem a la idea intuïtiva que havíem formulat al Capítol 3.

Tot el que hem fet fins aquí es generalitza sense dificultat a processos a valors en qualsevol espai de mesura. Però si a més tenim a l'espai d'estats una estructura topològica, la podem fer servir per escriure més equivalències. En particular, això valdrà per a processos a valors en $\mathbb{R}$ o en $\mathbb{R}^n$.

4.2.2 Proposició

Si $\{X_t,\ t \in \mathbb{R}^+\}$ és un procés a valors en un espai de mesura $(E, \mathfrak{E})$ tal que E és un espai polonès (mètric, separable i complet) i $\mathfrak{E}$ és al σ-àlgebra de Borel respecte aquesta topologia, aleshores X és Markov sii

$$\mathrm{E}\left[\psi(X_{t+h})/X_{s_1}, \dots, X_{s_n}, X_t\right] = \mathrm{E}\left[\psi(X_{t+h})/X_t\right] \quad ,$$

per a tota $\psi \colon E \to \mathbb{R}$ contínua i acotada.
I si, a més, E és localment compacte, aleshores és suficient prendre ψ contínua i a suport compacte. $\square$

Les formulacions anteriors de la propietat de Markov tenen un anàleg intercanviant passat per futur.

4.2.3 Exercici

Que un procés $\{X_t,\ t \in \mathbb{R}^+\}$ a valors en $\mathbb{R}^n$ sigui Markov no implica que les seves components ho siguin. Exemple:

$$\left\{\left(W_t, \int_0^t W_s\, ds\right),\ t \in \mathbb{R}^+\right\} \quad ,$$

on W és un procés de Wiener.

Al revés tampoc val: components Markov no implica vector Markov.

4.2.4 Exercici

Un procés $\{X_t,\ t \in \mathbb{R}^+\}$ es diu que té *increments independents del passat* si $\forall h > 0$, $X_{t+h} - X_t \perp\!\!\!\perp \{X_s,\ s \leqslant t\}$. ($X \perp\!\!\!\perp Y$ col dir que X i Y són independents.)
Un procés amb increments independents del passat és Markov.

4.3 Processos de Markov respecte una filtració

Una *filtració* $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ és una família creixent de σ-àlgebres.

Un procés estocàstic $\{X_t,\ t \in \mathbb{R}^+\}$ és adaptat a la filtració $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ si X_t és $\mathfrak{F}_t$-mesurable per a cada t.

4.3.1 Definició

Diem que un procés X és Markov respecte una filtració $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ si és adaptat i

$$\mathfrak{F}_t \underset{X_t}{\perp\!\!\!\perp} \sigma\{X_s,\ s \geqslant t\} \quad , \quad \forall t \in \mathbb{R}^+ \quad . \quad \square$$

Podem desenvolupar en aquest context totes les equivalències que hem fet abans. Per exemple, X és Markov respecte $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ sii $\forall t \in \mathbb{R}^+$, $\forall h > 0$,

$$P\left(X_{t+h} \in B/\mathfrak{F}_t\right) = P\left(X_{t+h} \in B/X_t\right) \quad , \quad \forall t \in \mathbb{R}^+,\ \forall B \in \mathfrak{B}(\mathbb{R}) \quad .$$

4.3.2 Observació

Si X és Markov respecte una filtració $\{\mathfrak{F}_t,\ t \in \mathbb{R}\}$, aleshores és Markov respecte la seva filtració natural $\mathfrak{F}_t = \sigma\{X_s,\ s \leqslant t\}$, o sigui en el sentit que hem treballat en els apartats anteriors. $\square$

4.4 Classes monòtones

Per demostrar les diverses simplificacions de la formulació de la propietat de Markov que hem vist, són útils els π-sistemes i els λ-sistemes, i un teorema de classes monòtones per a funcions.

Sigui Ω un conjunt i $\mathfrak{C}$ una col·lecció de subconjunts de Ω. Diem que $\mathfrak{C}$ és un *π-sistema* de $\mathcal{P}(\Omega)$ sii és tancat per interseccions finites. Diem que $\mathfrak{C}$ és un λ-sistema de $\mathcal{P}(\Omega)$ sii

a) $\Omega \in \mathfrak{C}$.

b) $A, B \in \mathfrak{C}$, $A \subset B \Rightarrow B - A \in \mathfrak{C}$.

c) $\{A_n\}_{n\in\mathbb{N}} \subset \mathfrak{C}$, $A_n \nearrow A \Rightarrow A \in \mathfrak{C}$.

Una col·lecció de conjunts és σ-àlgebra si i només si és a la vegada un π-sistema i un λ-sistema. Es pot demostrar que si $\mathfrak{C}$ és un π-sistema, aleshores el λ-sistema generat per $\mathfrak{C}$ coincideix amb la σ-àlgebra generada per $\mathfrak{C}$.

L'interès principal dels π-sistemes és que donen lloc al Teorema més fort possible sobre determinació de mesures: Si $\mathfrak{C}$ és π-sistema de $\mathcal{P}(\Omega)$, μ_1 i μ_2 són mesures finites sobre $(\Omega, \sigma(\mathfrak{C}))$, i $\mu_1 = \mu_2$ sobre $\mathfrak{C}$, aleshores $\mu_1 \equiv \mu_2$.

En la mateixa línia, el teorema següent, conegut com *Teorema de les Classes Monòtones (per funcions)* permet deduir resultats sobre funcions mesurables qualssevol a partir de resultats sobre els indicadors d'elements d'un π-sistema.

4.4.1 Teorema

Sigui $\mathfrak{C}$ un ϱpi-sistema de $\mathcal{P}(\Omega)$.
Sigui $\mathcal{H}$ una col·lecció de funcions reals acotades sobre Ω tal que

1) $\mathbf{1}_C \in \mathcal{H}$, $\forall C \in \mathfrak{C}$.

2) $1 \in \mathcal{H}$.

3) $\mathcal{H}$ és un espai vectorial.

4) Si $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$ és una successió no-decreixent de funcions no-negatives uniformement acotada, aleshores $f = \lim_{n\to\infty} f_n \in \mathcal{H}$.

Aleshores, $\mathcal{H}$ conté el conjunt de totes les funcions reals mesurables i acotades sobre $(\Omega, \sigma(\mathfrak{C}))$.

Demostració: Sigui $\mathfrak{D}$ la col·lecció de conjunts $A \subset \Omega$ tal que $\mathbf{1}_A \in \mathcal{H}$. És immediat de 2), 3), 4) que $\mathfrak{D}$ és un λ-sistema. Per 1), conté el π-sistema $\mathfrak{C}$. En conclusió, conté $\sigma(\mathfrak{C})$.

Sigui f una funció $\sigma(\mathfrak{C})$-mesurable i acotada. Suposem primer que f és positiva. Aleshores, és el límit puntual d'una successió creixent $\{f_n\}_{n\in\mathbb{N}}$ de funcions elementals $\sigma(\mathfrak{C})$-mesurables, és a dir, de combinacions lineals d'indicadors d'elements de $\sigma(\mathfrak{C})$. Sabem que aquests indicadors són de $\mathcal{H}$. Per 3), $f_n \in \mathcal{H}$, $\forall n \in \mathbb{N}$. Per 4), $\lim_{n\to\infty} f_n = f \in \mathcal{H}$.

Si f és $\sigma(\mathfrak{C})$-mesurable i acotada, aleshores f^+ i f^- també i a més són positives; pel cas anterior, $f^+, f^- \in \mathcal{H}$. Per 3), $f \in \mathcal{H}$.

5. Funcions de transició

5.1 Teorema de la Mesura Producte i lleis condicionades

Recordem el Teorema de la Mesura Producte, especialitzat al cas de probabilitats (per una versió per a mesures més generals, i la demostració, vegeu Alabert [1]).

5.1.1 Teorema de la Mesura Producte

Sigui $(\Omega_1, \mathfrak{F}_1, \mu)$ *un espai de probabilitat.*
Sigui $(\Omega_2, \mathfrak{F}_2)$ *un espai mesurable.*
Sigui $\tau\colon \Omega_1 \times \mathfrak{F}_2 \longrightarrow [0, +\infty]$ *una probabilitat de transició.*
Aleshores:

1) Existeix una única probabilitat ν *en* $(\Omega_1 \times \Omega_2, \mathfrak{F}_1 \otimes \mathfrak{F}_2)$ *tal que*

$$\nu(A_1 \times A_2) = \int_{A_1} \tau(\omega_1, A_2)\, \mu(d\omega_1) \quad .$$

2) Si $X\colon \Omega_1 \times \Omega_2 \longrightarrow \bar{\mathbb{R}}$ *és una funció* $(\mathfrak{F}_1 \otimes \mathfrak{F}_2, \mathfrak{B}(\bar{\mathbb{R}}))$*-mesurable i la integral de* X *respecte* ν *existeix, aleshores:*
La funció

$$\begin{aligned} \Omega_1 &\longrightarrow \bar{\mathbb{R}} \\ \omega_1 &\longmapsto \int_{\Omega_2} X(\omega_1, \omega_2)\, \tau(\omega_1, d\omega_2) \end{aligned}$$

està ben definida μ*-q.p.t.* $\omega_1 \in \Omega_1$*, és mesurable (definida com a zero, per exemple, allà on la integral no existeixi), la seva integral respecte* μ *existeix i*

$$\int_{\Omega_1 \times \Omega_2} X\, d\nu = \int_{\Omega_1} \Big(\int_{\Omega_2} X(\omega_1, \omega_2)\, \tau(\omega_1, d\omega_2) \Big)\, \mu(d\omega_1) \quad .$$

El condicionament segueix una direcció contrària al del Teorema de la Mesura Producte. Una probabilitat en un espai producte es desintegra en un mesura inicial i una probabilitat de transició.

5.1.2 Definició

Siguin $X, Y\colon \Omega \longrightarrow \mathbb{R}$ variables aleatòries.
S'anomena *llei de Y condicionada per X* a tota probabilitat de transició p de $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ en $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ que compleixi

$$P(\{X \in A, Y \in B\}) = \int_A p(x, B)\, P_X(dx)\ , \quad \forall A, B \in \mathfrak{B}(\mathbb{R}) \quad . \quad \square$$

De vegades s'escriu $p(x, B) =: P\{Y \in B / X = x\}$ Però cal anar amb compte amb aquesta notació, perquè no es tracta d'un nombre que existeixi per a cada x; és una funció que està definida quasi per tot respecte la llei de X. Sempre existeix la llei d'una variable condicionada a una altra (vegeu el Capítol 2). Es té també que, si p_1 i p_2 són dues d'elles, aleshores $p_1(x, B) = p_2(x, B)$, $\forall B \in \mathfrak{B}(\mathbb{R})$, P_X-q.s. Dit d'una altre manera, la llei del vector (X, Y) en $\mathbb{R}^2$ es pot desintegrar en la llei de X i una probabilitat de transició de $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ en $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.

5.2 Funcions de transició

Pensem ara en les variables X_s i X_t d'un procés estocàstic, $s < t$, i en la llei de X_t condicionada per X_s. Afegim als arguments de la llei condicionada (probabilitat de transició) els índexos s i t, per recordar de quines variables es tracta:

$$P(\{X_s \in A, X_t \in B\}) = \int_A p(s, x; t, B)\, P_{X_s}(dx)\ , \quad \forall A, B \in \mathfrak{B}(\mathbb{R}) \quad .$$

Per tant, podem pensar en p com una funció de quatre arguments que compleix:

1) $p(s, \cdot; t, B)$ és mesurable, $\forall s \leqslant t \in \mathbb{R}^+$, $\forall B \in \mathfrak{B}(\mathbb{R})$.

2) $p(s, x; t, \cdot)$ és una probabilitat sobre $\mathfrak{B}(\mathbb{R})$, $\forall s \leqslant t \in \mathbb{R}^+$, $\forall x \in \mathbb{R}$.

A més, està clar que ha de ser

3) $p(s, x; s, \cdot) = \delta_x$, $\forall s \in \mathbb{R}^+$, $\forall x \in \mathbb{R}$.

Una funció p complint 1), 2) i 3) s'anomena *funció de transició*.

Aquesta funció p existeix per a tot procés estocàstic real. Noteu que, donat un procés, el valor $p(s, x; t, B)$ està determinat llevat de conjunts B amb $P_{X_s}(B) = 0$.

La funció de transició no determina la llei del procés. Si donem la llei de la primera variable X_0, la funció de transició ens determinarà la llei de cada X_t i les lleis bidimensionals (X_s, X_t), gràcies al Teorema de la Mesura Producte, però això tampoc no és suficient per determinar la llei del procés, en general. En canvi, si el procés és Markov, la funció p i la llei de la primera variable X_0, determinen la llei del procés:

5.2.1 Proposició

Si $\{X_t,\ t \in \mathbb{R}^+\}$ és Markov respecte $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$, la seva funció de transició més la llei de X_0 determinen la llei del procés.

Demostració:

Siguin $s < t_1 < \cdots < t_n \in \mathbb{R}^+$.

Siguin $\varphi_1, \ldots, \varphi_n$ funcions mesurables fitades.

$$
\begin{aligned}
&\mathrm{E}\left[\varphi_1(X_{t_1}) \cdot \ldots \cdot \varphi_n(X_{t_n})\right] = \\
&\mathrm{E}\left[\,\mathrm{E}\left[\varphi_1(X_{t_1}) \cdot \ldots \cdot \varphi_n(X_{t_n})/X_s\right]\right] = \\
&\mathrm{E}\left[\,\mathrm{E}\left[\varphi_1(X_{t_1}) \cdot \ldots \cdot \varphi_n(X_{t_n})/\mathfrak{F}_s\right]\right] = \\
&\mathrm{E}\left[\,\mathrm{E}\left[\mathrm{E}\left[\varphi_1(X_{t_1}) \cdot \ldots \cdot \varphi_n(X_{t_n})/\mathfrak{F}_{t_{n-1}}\right]/\mathfrak{F}_s\right]\right] = \\
&\mathrm{E}\left[\,\mathrm{E}\left[\varphi_1(X_{t_1}) \cdot \ldots \cdot \varphi_{n-1}(X_{t_{n-1}})\,\mathrm{E}\left[\varphi_n(X_{t_n})/\mathfrak{F}_{t_{n-1}}\right]/\mathfrak{F}_s\right]\right] = \\
&\mathrm{E}\left[\,\mathrm{E}\left[\varphi_1(X_{t_1}) \cdot \ldots \cdot \varphi_{n-1}(X_{t_{n-1}})\,\mathrm{E}\left[\varphi_n(X_{t_n})/X_{t_{n-1}}\right]/\mathfrak{F}_s\right]\right] = \\
&\mathrm{E}\left[\,\mathrm{E}\left[\varphi_1(X_{t_1}) \cdot \ldots \cdot \varphi_{n-1}(X_{t_{n-1}}) \int_{\mathbb{R}} \varphi_n(y_n)\, p(t_{n-1}, X_{t_{n-1}}; t_n, dy_n)/\mathfrak{F}_s\right]\right] = \\
&\ldots\ \ldots
\end{aligned}
$$

Condicionem successivament per $\mathfrak{F}_{t_{n-2}}, \ldots, \mathfrak{F}_{t_1}$, i obtenim

$$
\begin{aligned}
&\mathrm{E}\Big[\int_{\mathbb{R}} \varphi_1(y_1) \int_{\mathbb{R}} \varphi_2(y_2) \cdots \int_{\mathbb{R}} \varphi_n(y_n) \\
&\quad p(t_{n-1}, y_{n-1}; t_n, dy_n) p(t_{n-2}, y_{n-2}; t_{n-1}, dy_{n-1}) \cdot \ldots \cdot p(s, X_s; t_1, dy_1)\Big] = \\
&\int_{\mathbb{R}} \int_{\mathbb{R}} \varphi_1(y_1) \int_{\mathbb{R}} \varphi_2(y_2) \cdots \int_{\mathbb{R}} \varphi_n(y_n) \\
&\quad p(t_{n-1}, y_{n-1}; t_n, dy_n) p(t_{n-2}, y_{n-2}; t_{n-1}, dy_{n-1}) \cdot \ldots \cdot p(s, x; t_1, dy_1) P_{X_s}(dx) \quad .
\end{aligned}
$$

Particularitzem a φ indicadors i $s = 0$:

$$
\begin{aligned}
&P\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\} = \\
&\int_{\mathbb{R}} \int_{B_1} \int_{B_2} \cdots \int_{B_n} p(t_{n-1}, y_{n-1}; t_n, dy_n) p(t_{n-2}, y_{n-2}; t_{n-1}, dy_{n-1}) \cdot \ldots \cdot p(s, x; t_1, dy_1) P_{X_0}(dx) \quad .
\end{aligned}
$$

Queda clar que tot depèn de la lei de X_0 i la funció de transició. □

Observem el següent cas particular del càlcul anterior: Siguin $s < r < t$. Calcularem $P\big(X_t \in B/\mathfrak{F}_s\big)$ de dues maneres.

Per una banda,

$$
P\big(X_t \in B/\mathfrak{F}_s\big) = \int_B p(s, X_s; t, dy) = p(s, X_s; t, B) \quad . \tag{5.2.1}
$$

Per altra banda,

$$
\begin{aligned}
P\big(X_t \in B/\mathfrak{F}_s\big) &= P\big(X_t \in B,\ X_r \in \mathbb{R}/\mathfrak{F}_s\big) \\
&= \int_{\mathbb{R}} \int_B p(r, y; t, dz) \cdot p(s, X_s; r, dy) = \int_{\mathbb{R}} p(r, y; t, B) \cdot p(s, X_s; r, dy)
\end{aligned}
$$

Hem obtingut el

5.2.2 Corol·lari

Si p és la funció de transició d'un procés de Markov $\{X_t,\ t \in \mathbb{R}\}$ respecte una filtració $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$, aleshores, $\forall s < r < t \in \mathbb{R}^+$,

$$
p(s, X_s; t, B) = \int_{\mathbb{R}} p(r, y; t, B) \cdot p(s, X_s; r, dy) \quad .
$$

Aquesta igualtat és (una versió de) l'equació de Chapman–Kolmogorov.

5.2.3 Corol·lari

En la situació anterior,

$$p(s,x;t,B) = \int_{\mathbb{R}} p(r,y;t,B)\cdot p(s,x;r,dy) \quad , \quad P_{X_s}\text{-q.s. } x\in\mathbb{R} \quad . \tag{5.2.2}$$

(Una altra versió, més habitual, de l'equació de Chapman–Kolmogorov).
(Recordeu l'Exercici 3.2.2. Allà el P_X-q.s. no és necessari, per la forma explícita que tenim de la densitat).

5.2.4 Definició

Una funció de transició que compleixi l'equació de Chapman–Kolmogorov s'anomena *funció de transició markoviana.* □

Tot el que hem fet val per a processos a valors en $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$ i, de fet, per espais més generals encara (polonesos, per exemple).

Està clar que la funció de transició d'un procés de Markov és una funció de transició markoviana (és el que estem afirmant en els corol·laris anteriors). I també és important recordar que compleix

$$P\big(X_t\in B/_{\mathfrak{F}_s}\big) = p(s,X_s;t,B)$$

(deduït en (5.2.1).)

Hi ha dues qüestions naturals més sobre funcions de transició markovianes i processos de Markov:

1. Un procés que tingui una funció de transició markoviana, és Markov (almenys respecte la seva filtració natural)? Resposta: No necessàriament.
2. Donada una funció de transició markoviana, existeix algun procés de Markov que la té per funció de transició? Resposta: Sí. (Teorema següent). I, a més, serà únic en llei si fixem la llei de X_0, segons hem vist abans.

5.2.5 Teorema

Sigui $p(s,x;t,B)$ una funció de transició markoviana.
Sigui μ una probabilitat en $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$.
Definim una família de lleis en dimensió finita mitjançant:

$$P_0(B) = \mu(B) \quad , \quad B\in\mathfrak{B}(\mathbb{R}^d) \quad ,$$

i, per $0\leqslant t_1 < \cdots < t_{n+1}$,

$$P_{t_1,\dots,t_{n+1}}(B) = \int_B p(t_n,y_n;t_{n+1},dy_{n+1})\, P_{t_1,\dots,t_n}(dy_1,\dots,dy_n) \quad , \quad B\in\mathfrak{B}\big((\mathbb{R}^d)^{n+1}\big) \quad .$$

Aleshores, $P_{t_1,\dots,t_n}$ és una família consistent de lleis en dimensió finita, i determinen una probabilitat P en $\big((\mathbb{R}^d)^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}\big)$ tal que el procés coordenat $X_t(\omega)=\omega(t)$ és un procés de Markov amb funció de transició p i llei inicial μ.

Demostració:
(1). (Consistència).
És bastant evident, però fem-ho:

Hem de veure que $P_K = P_J \circ \pi_{JK}^{-1}$, on $K = \{t_1, \ldots, t_k\}$, $J = \{t_1, \ldots, t_k, t_{k+1}, \ldots, t_j\}$, i π_{JK} és la projecció a les primeres k components de $\mathbb{R}^j$.

Sigui $B \in \mathfrak{B}((\mathbb{R}^d)^k)$.

$$P_{t_1,\ldots,t_j}(\pi_{JK}^{-1}(B)) = P_{t_1,\ldots,t_j}(B \times \mathbb{R}^{j-k})$$
$$= \int_{B\times\mathbb{R}^{j-k}} p(t_{j-1}, y_{j-1}; t_j, dy_j)\, P_{t_1,\ldots,t_{j-1}}(dy_1, \ldots, dy_{j-1}) = \cdots$$
$$= \int_{B\times\mathbb{R}^{j-k}} p(t_{j-1}, y_{j-1}; t_j, dy_j) \cdot \ldots \cdot p(t_k, y_k; t_{k+1}, dy_{k+1})\, P_{t_1,\ldots,t_k}(dy_1, \ldots, dy_k) \quad .$$

Aplicant diverses vegades l'equació de Chapman–Kolmogorov (5.2.2),

$$\int_{\mathbb{R}^{j-k}} p(t_{j-1}, y_{j-1}; t_j, dy_j) \cdot \ldots \cdot p(t_k, y_k; t_{k+1}, dy_{k+1})$$
$$= \int_{\mathbb{R}^{j-k-1}} p(t_{j-1}, y_{j-1}; t_j, dy_j) \cdot \ldots \cdot p(t_{k-1}, y_{k-1}; t_k, dy_k) = \cdots$$
$$= \int_{\mathbb{R}} p(t_{j-1}, y_{j-1}; t_j, dy_j) = 1 \quad .$$

Per tant, queda

$$P_{t_1,\ldots,t_j}(\pi_{JK}^{-1}(B)) = \int_B P_{t_1,\ldots,t_k}(dy_1, \ldots, dy_k) = P_{t_1,\ldots,t_k}(B) \quad .$$

(2). (El que queda).

Siguin $0 \leqslant r_1 < \cdots < r_n = s < t$, $B_1, \ldots, B_n, B \in \mathfrak{B}(\mathbb{R}^d)$.

$$P\{X_{r_1} \in B_1, \ldots, X_{r_n} \in B_n, X_t \in B\} = P_{r_1,\ldots,r_n,t}(B_1 \times \cdots \times B_n \times B)$$
$$= \int_{B_1\times\cdots\times B_n\times B} p(s, y_n; t, dy_{n+1})\, P_{r_1,\ldots,r_n}(dy_1, \ldots, dy_n)$$
$$= \int_{\{X_{r_1}\in B_1,\ldots,X_{r_n}\in B_n\}} p(s, X_s(\omega); t, B)\, P(d\omega) \quad ,$$

això últim pel Teorema de la Mesura Imatge. Això ens diu (vegeu la definició de probabilitat condicionada) que

$$(\omega, B) \longmapsto p(s, X_s(\omega); t, B)$$

és la probabilitat condicionada de X_t a $\sigma\{X_{r_1}, \ldots, X_{r_n}\}$. És a dir,

$$p(s, X_s(\omega); t, B) = P\big(X_t \in B / X_{r_1}, \ldots, X_{r_n}\big) \quad ,$$

d'on X és un procés de Markov amb funció de transició p:

$$P\big(X_t \in B / X_s\big) = P\big(X_t \in B / X_{r_1}, \ldots, X_{r_n}\big)$$

(recordem que $r_n = s$).

5.2.6 Exercicis

1) Si $\{X_t,\ t \in \mathbb{R}^+\}$ és un procés amb increments independents del passat, aleshores és Markov, i la seva funció de transició compleix
$$p(s, x; t, B) = p(s, x + h; t, B + h) \quad , \quad \forall h > 0 \quad .$$
Una funció de transició amb aquesta propietat es diu que és *espacialment homogènia*. El procés de Markov que determina és *espacialment homogeni*. De fet, si un procés de Markov és espacialment homogeni, aleshores té increments independents.

2) Si W és un procés de Wiener estàndar començant en 0, el procés $\{|W_t|,\ t \in \mathbb{R}^+\}$ s'anomena *procés de Wiener amb reflexió en zero*.
És un procés de Markov. Quina és la funció de transició?

3) Si W és un procés de Wiener estàndar començant en 0, la seva part positiva, $\{W_t^+,\ t \in \mathbb{R}^+\}$, no és Markov. De tota manera, quina és la seva funció de transició?

6. Família markoviana

Hem vist que la llei d'un procés de Markov queda determinada per la funció de transició i la llei de la primera variable. Per tant, si μ és una probabilitat en $\mathbb{R}$ (o $\mathbb{R}^d$, si el procés pren valors en $\mathbb{R}^d$) un *procés de Markov amb llei inicial* μ en l'espai de probabilitat $(\Omega, \mathfrak{F}, P)$ respecte a la filtració $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ voldrà dir un procés estocàstic $\{X_t,\ t \in \mathbb{R}^+\}$ tal que

a) $P\big(X_t \in B /_{\mathfrak{F}_s}\big) = P\big(X_t \in B /_{X_s}\big)$, q.s., $\forall t \geqslant s$, $\forall B \in \mathfrak{B}(\mathbb{R}^d)$.

b) $P\{X_0 \in B\} = \mu(B)$, $\forall B \in \mathfrak{B}(\mathbb{R}^d)$.

Al estudiar processos de Markov és interessant considerar a la vegada una família de lleis inicials en lloc d'una sola. Concretament, interessa considerar a la vegada totes les lleis δ_x, Delta de Dirac en $x \in \mathbb{R}^d$. Això ens porta al concepte de família markoviana. Hi ha diverses maneres de definir-lo.

6.1 Concepte de família markoviana

6.1.1 Definició

Sigui Ω un conjunt (de moment, no hi posem cap σ-àlgebra ni cap probabilitat).
Sigui, per a cada $t \in \mathbb{R}^+$, una aplicació $X_t \colon \Omega \to \mathbb{R}^d$.
Sigui $p(s, x; t, B)$ una funció de transició:

1) $p(s, \cdot; t, B)$ és mesurable, $\forall s \leqslant t \in \mathbb{R}^+$, $\forall B \in \mathfrak{B}(\mathbb{R}^d)$.

2) $p(s, x; t, \cdot)$ és una probabilitat en $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$, $\forall s \leqslant t$, $\forall x \in \mathbb{R}^d$.

3) $p(s, x; s, \cdot) = \delta_x$, $\forall s \in \mathbb{R}^+$, $\forall x \in \mathbb{R}^d$.

Per a tot $s \in \mathbb{R}^+$, i per a tot $x \in \mathbb{R}^d$, sigui $P_{s,x}$ una probabilitat sobre la σ-àlgebra de Ω generada per $\{X_t,\ t \geqslant s\}$, que denotarem $\mathfrak{F}^s(X)$.
Diem que $\{X_t,\ t \in \mathbb{R}^+\}$, $\{P_{s,x}\}_{s \in \mathbb{R}^+, x \in \mathbb{R}^d}$ és una *família markoviana* amb funció de transició p si

1) $\forall s \in \mathbb{R}^+$, $\forall x \in \mathbb{R}^d$, la família d'aplicacions $\{X_t,\ t \geqslant s\}$ és un procés de Markov a l'espai de probabilitat $(\Omega, \mathfrak{F}^s(X), P_{s,x})$ amb funció de transició p.

2) $P_{s,x}\{X_s = x\} = 1$. □

6.1.2 Observacions

1) La definició val igual per processos amb espai d'estats qualsevol.
2) Si no es vol considerar processos que comencen en instants de temps estrictament positius s, es poden definir com a constants abans de s.
3) Hi ha definicions més generals tant de procés de Markov com de família markoviana, que permeten que les trajectòries "es morin" en instants de temps que depenen de cada trajectòria. Aquests instants s'anomenen *temps de vida* de la trajectòria. Així ho fa per exemple Dynkin [3]. □

6.1.3 Proposició

En la situació de la definició anterior:
Siguin $s \leqslant t_1 < \cdots < t_n \in \mathbb{R}^+$.
Aleshores:

$$P_{s,x}\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\} =$$
$$\int_{B_1}\int_{B_2}\cdots\int_{B_n} p(t_{n-1}, y_{n-1}; t_n, dy_n)p(t_{n-2}, y_{n-2}; t_{n-1}, dy_{n-1}) \cdot \ldots \cdot p(s, x; t_1, dy_1) \quad .$$

Demostració: Es fa semblantment a la de la Proposició 5.2.1. Observeu que la integral més exterior d'aquella demostració aquí està feta, aprofitant que la llei de X_s sota la probabilitat $P_{s,x}$ és la Delta de Dirac en x.

6.1.4 Corol·lari 1

1) Si $u \leqslant s \leqslant t_1 < \cdots < t_n$, *aleshores*
$$P_{u,x}\big(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n /_{\mathfrak{F}_s(X)}\big) = P_{s,X_s}\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\} \quad , \quad q.s.$$
2) Si $s \leqslant t$,
$$P_{s,x}\{X_t \in B\} = p(s, x; t, B) \quad .$$

Demostració: Vegeu altre cop la demostració de la Proposició 5.2.1. Per calcular la probabilitat condicionada de 1), la integral més exterior no s'ha de fer; per tant el diferencial més exterior és $p(s, X_s; t_1, dy)$, i obtenim la fórmula de
$$P_{s,X_s}\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\} \quad .$$

La fórmula 2) és simplement un cas particular de l'enunciada a la proposició anterior.

6.1.5 Corol·lari 2

En la situació de la proposició anterior, si $s < r < t \in \mathbb{R}^+$,
$$p(s, x; t, B) = \int_{\mathbb{R}^d} p(r, y; t, B) \cdot p(s, x; r, dy) \quad .$$
(Equació de Chapman–Kolmogorov.)

Demostració:
$$p(s, x; t, B) = P_{s,x}\{X_t \in B\} = P_{s,x}\{X_t \in B,\ X_r \in \mathbb{R}^d\} = \int_{\mathbb{R}^d} p(r, y; t, B) \cdot p(s, x; r, dy) \quad .$$
Observeu que aquí no cal posar el q.s. respecte la llei de X_s que posavem en el Corol·lari 5.2.3. □

Hem vist doncs que la funció de transició d'una família markoviana satisfà 1), 2), 3) de la Definició 6.1.1, i

4) L'equació de Chapman–Kolmogorov del Corol·lari 6.1.5.

Resulta que donada una funció complint 1), 2), 3), 4), existeix una família markoviana que la té per funció de transició:

6.1.6 Teorema

Sigui p una funció de transició complint l'equació de Chapman–Kolmogorov del Corol·lari 6.1.5.
Aleshores existeix una família markoviana amb funció de transició p.
(Val l'enunciat no sols per $\mathbb{R}^d$ sino per qualsevol espai polonès amb la σ-àlgebra boreliana com a espai d'estats.)

Demostració: Es fa via el Teorema de Daniell–Kolmogorov:

1) Per a cada s i x, es construeix l'espai canònic: $(\mathbb{R}^d)^{[s,\infty[}$ amb la σ-àlgebra producte i una probabilitat que notem $\tilde{P}_{s,x}$.
2) Prenem el conjunt $\Omega = (\mathbb{R}^d)^{\mathbb{R}^+}$, amb la σ-àlgebra producte; definim les probabilitats $P_{s,x}$ sobre aquest espai mesurable de la manera següent:

$$P_{s,x}(A) = \tilde{P}_{s,x}(\tilde{A}) \quad , \quad \forall A \in \mathfrak{F}^s(X) \quad ,$$

on

$$\tilde{A} := \{\tilde{\omega} \in (\mathbb{R}^d)^{[s,\infty[} : \ \exists \omega \in A \text{ tal que } \omega = \tilde{\omega} \text{ en } [s,\infty[\} \quad .$$

3) Definim $X_t(\omega) = \omega(t)$.

6.1.7 Proposició

Sigui $\{X_t, \ t \in \mathbb{R}^+\}$, $\{P_{s,x}\}_{s\in\mathbb{R}^+, x\in\mathbb{R}^d}$, una família markoviana.
Sigui μ una probabilitat en $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$.
Definim en $(\Omega, \mathfrak{F}(X))$ la probabilitat

$$P(A) := \int_{\mathbb{R}^d} P_{0,x}(A)\, \mu(dx) \quad , \quad A \in \mathfrak{F}(X) \quad . \tag{6.1.1}$$

Aleshores, $\{X_t, \ t \in \mathbb{R}\}$ és un procés de Markov en $(\Omega, \mathfrak{F}(X), P)$ amb distribució inicial μ i la mateixa funció de transició de la família markoviana.

Demostració: Sabem que $\{X_t, \ t \in \mathbb{R}^+\}$ és un procés de Markov a $(\Omega, \mathfrak{F}(X), P_{0,x})$ amb la funció de transició p de la família markoviana. Sabem també que $P_{0,x}\{X_0 = x\} = 1$.
La funció $x \mapsto P_{0,x}(A)$ és mesurable per a tot $A \in \mathfrak{F}(X)$: En efecte, si A és de la forma $\{X_{t_1} \in B_1, \dots, X_{t_n} \in B\}$ això es desprèn de la fórmula de la Proposició 6.1.3. Aquests conjunts formen un π-sistema que genera $\mathfrak{F}(X)$. És trivial comprovar que els elements del λ-sistema que genera aquest π-sistema compleixen la propietat de mesurabilitat que es busca. Per tant, la mesurabilitat de $x \mapsto P_{0,x}(A)$ és vàlida $\forall A \in \mathfrak{F}(X)$.
Finalment, amb P definida per (6.1.1),

$$P\{X_0 \in B\} = \int_{\mathbb{R}^d} P_{0,x}\{X_0 \in B\}\, \mu(dx) = \int_{\mathbb{R}^d} \mathbf{1}_B(x)\, \mu(dx) = \mu(B) \quad . \quad \square$$

6.2 Exemples de famílies markovianes

6.2.1 Exemple (família de Wiener)

Una *família de Wiener d-dimensional* (o *família browniana*) és un procés $W \equiv \{W_t,\ t \in \mathbb{R}^+\}$ sobre un espai mesurable $(\Omega, \mathfrak{F})$, juntament amb una família de probabilitats $\{P_{s,x}\}_{s\in\mathbb{R}^+, x\in\mathbb{R}^d}$ tals que

1) $\forall A \in \mathfrak{F}(W)$, $\forall s \in \mathbb{R}^+$, l'aplicació $x \mapsto P_{s,x}(A)$ és mesurable.
2) $\forall x \in \mathbb{R}^d$, $P_{s,x}\{W_s = x\} = 1$.
3) $\forall s \in \mathbb{R}^+$, $\forall x \in \mathbb{R}^d$, el procés $\{W_t,\ t \geqslant s\}$ és, sota $P_{s,x}$, un procés de Wiener d-dimensional començant en x.

6.2.2 Exemple (família de Wiener amb dispersió a i deriva b)

Suposem que $\{W_t,\ t \in \mathbb{R}^+\}$, $\{P_{s,x}\}_{s\in\mathbb{R}^+, x\in\mathbb{R}^d}$, constitueixen una família de Wiener.
Sigui a una matriu $d \times d$ no singular i sigui $b \in \mathbb{R}^d$.
Aleshores, el procés $X_t = aW_t + bt$ sobre $(\Omega, \mathfrak{F})$, amb la família de probabilitats $\{P_{s,a^{-1}x}\}$ s'anomena *família de Wiener amb dispersió a i deriva b.*
Nota: A la deriva tothom l'anomena *drift*.

6.2.3 Exemple (família de Poisson)

Una *família de Poisson amb intensitat* $\lambda > 0$ és un procés $N \equiv \{N_t,\ t \in \mathbb{R}^+\}$ sobre un espai mesurable $(\Omega, \mathfrak{F})$, juntament amb una família de probabilitats $\{P_{s,x}\}_{s\in\mathbb{R}^+, x\in\mathbb{R}^d}$ tals que

1) $\forall A \in \mathfrak{F}(N)$, $\forall s \in \mathbb{R}^+$, l'aplicació $x \mapsto P_{s,x}(A)$ és mesurable.
2) $\forall x \in \mathbb{R}^d$, $P_{s,x}\{N_s = x\} = 1$.
3) $\forall s \in \mathbb{R}^+$, $\forall x \in \mathbb{R}^d$, el procés $\{N_t - N_s,\ t \geqslant s\}$ és, sota $P_{s,x}$, un procés de Poisson amb intensitat λ.

6.2.4 Exercici

Els exemples anteriors són famílies markovianes.

6.3 Operadors shift

Sigui Ω un conjunt i sigui, per a cada $t \in \mathbb{R}^+$ una aplicació $X_t\colon \Omega \to \mathbb{R}^d$.

Suposem que $\forall \omega \in \Omega$, $\forall h \in \mathbb{R}^+$, existeix un únic $\omega' \in \Omega$ tal que $X_t(\omega') = X_{t+h}(\omega)$, $\forall t \in \mathbb{R}^+$.

Sigui $\theta_h\colon \Omega \to \Omega$ l'aplicació que relaciona ω i ω': $\theta_h\omega = \omega'$. L'operador θ_h fa un "shift" de les trajectòries de $\{X_t,\ t \in \mathbb{R}^+\}$ una distància h cap a l'esquerra.

Per a cada $A \subset \Omega$, podem considerar el conjunt $\theta_h^{-1}A = \{\omega \in \Omega :\ \theta_h\omega \in A\}$. Per exemple,

a) Si $A = \{X_t \in B\}$,

$$\begin{aligned}\theta_h^{-1}A &= \{\omega :\ \theta_h\omega \in A\} = \{\omega :\ \theta_h\omega \in \{X_t \in B\}\} \\ &= \{\omega :\ X_t(\theta_h\omega) \in B\} = \{\omega :\ X_{t+h}(\omega) \in B\} = \{X_{t+h} \in B\} \quad .\end{aligned}$$

b) Si $A = \{\lim_{t\to\infty} X_t = 0\}$,

$$\begin{aligned}\theta_h^{-1}A &= \{\omega :\ \theta_h\omega \in A\} = \{\omega :\ \lim_{t\to\infty} X_t(\theta_h\omega) = 0\} \\ &= \{\omega :\ \lim_{t\to\infty} X_{t+h}(\omega) = 0\} = A \quad .\end{aligned}$$

Sigui ara η una funció sobre Ω, valuada no importa on. Definim el "shift" de η com

$$[\theta_h \eta](\omega) = \eta(\theta_h \omega) \quad .$$

(Observeu que estem usant la notació θ_h per dues coses diferents.)

Per exemple,

a) Si $\eta = X_t$,

$$[\theta_h X_t](\omega) = X_t(\theta_h \omega) = X_{t+h}(\omega) \quad .$$

(naturalment! Els shifts estan construïts per tal que fagin això sobre X_t.)

b) Si $\eta(\omega) = \int_0^t X_s(\omega)\,ds$,

$$\begin{aligned}[\theta_h \eta](\omega) = \eta(\theta_h \omega) &= \int_0^t X_s(\theta_h \omega)\,ds \\ = \int_0^t X_{s+h}(\omega)\,ds &= \int_h^{t+h} X_s(\omega)\,ds \quad .\end{aligned}$$

c) Si $\eta(\omega) = \inf\{t:\ X_t(\omega) \in B\}$,

$$\begin{aligned}[\theta_h \eta](\omega) &= \eta(\theta_h \omega) = \inf\{t:\ X_t(\theta_h \omega) \in B\} \\ &= \inf\{t:\ X_{t+h}(\omega) \in B\} = \inf\{t:\ t \geqslant h,\ X_t(\omega) \in B\} - h \quad .\end{aligned}$$

Suposem ara que $\{X_t, t \in \mathbb{R}^+\}$ és un procés estocàstic en un cert espai de probabilitat $(\Omega, \mathfrak{F}, P)$.

Considerem les σ-àlgebres generades pel procés: Notarem $\mathfrak{F}_{[s,t]}(X) := \sigma\{X_r,\ s \leqslant r \leqslant t\}$.

Si $A = \{X_r \in B\}$, $B \in \mathfrak{B}(\mathbb{R}^d)$, sabem que $\theta_h^{-1} A = \{X_{r+h} \in B\}$. Per tant, la σ-àlgebra generada per

$$\{\theta_h^{-1} A:\ A \in \mathfrak{F}_{[s,t]}(X)\}$$

és la generada per $\{X_{r+h}:\ s \leqslant r \leqslant t\}$, que és

$$\sigma\{X_r:\ s+h \leqslant r \leqslant t+h\} = \mathfrak{F}_{[s+h,t+h]}(X) \quad .$$

En particular, la σ-àlgebra $\mathfrak{F}(X)$ es transforma en $\mathfrak{F}^h(X)$.

Tornant al principi d'aquest apartat, observeu que a l'espai canònic es compleix la existència i unicitat del ω' tal que $X_t(\omega') = X_{t+h}(\omega)$, perquè els esdeveniments elementals es corresponen bijectivament amb les trajectòries del procés. En general, si no hi ha unicitat, els shifts no estan ben definits i no podrem fer $\theta_h^{-1} A$. Però no hi ha problema per conjunts $A \in \mathfrak{F}(X)$:

Si ω'_1 i ω'_2 compleixen, aleshores $\forall A \in \mathfrak{F}(X)$, o bé $\omega'_1, \omega'_2 \in A$ o bé $\omega'_1, \omega'_2 \notin A$. En efecte: Sigui $\mathfrak{C}$ la col·lecció de conjunts tals que passa això. Llavors, $\mathfrak{C}$ és una σ-àlgebra que conté els esdeveniments $\{X_t \in B\}$; per tant $\mathfrak{C} \supset \mathfrak{F}(X)$. Això fa que al definir $\theta_h^{-1} A = \{\omega:\ \omega' \in A\}$, no importa quin ω' s'agafi ($\theta_h \omega$ es defineix com un qualsevol d'ells).

7. Processos de Markov (temporalment) homogenis

7.1 Conceptes d'estacionarietat i homogeneïtat

7.1.1 Definició

Un procés $\{X_t,\ t \in \mathbb{R}^+\}$ és *estacionari* si les seves lleis en dimensió finita són invariants per trasllacions temporals, és a dir, la llei de $(X_{t_1}, \ldots, X_{t_n})$ coincideix amb la de $(X_{t_1+h}, \ldots, X_{t_n+h})$, per a tot $h > 0$.

7.1.2 Definició

Un procés $\{X_t,\ t \in \mathbb{R}^+\}$ té *increments estacionaris* si la llei de la variable aleatòria $X_t - X_s$ depèn només de la diferència $t - s$. □

Està clar que un procés estacionari té increments estacionaris.

Hi ha també el concepte de *procés estacionari en sentit ampli*, que normalment es refereix a processos de quadrat integrable les variables del qual tenen esperança constant i covariància que depèn només de la distància entre les variables a l'eix temporal. De vegades els processos estacionaris s'anomenen "estrictament estacionaris", per marcar més la diferència.

Un altre nom per als processos amb increments estacionaris és *processos homogenis*, o encara *temporalment homogenis*, per no confondre amb la homogeneïtat espacial, que és una altra cosa. Farem servir aquest terme, que és més curt.

Per exemple: El procés de Wiener té increments estacionaris, però no és estacionari. El procés de Ornstein–Uhlenbeck és estacionari.

7.2 Funcions de transició homogènies

La funció de transició d'un procés de Markov homogeni ($\equiv$ amb increments estacionaris) compleix

$$p(s, x; t, B) = p(s + h, x; t + h, B) \quad ,$$

i per tant ens podem estalviar un argument. Podem escriure

$$p(x; t, B) := p(0, x; t, B) = P\big(X_t \in B / X_0 = x\big) = P\big(X_{t+h} \in B / X_h = x\big) \quad .$$

Es diu que p és una *funció de transició homogènia.* $\{p(\cdot;t,\cdot),\ t\in\mathbb{R}^+\}$ és una família uniparamètrica de probabilitats de transició.

L'equació de Chapman–Kolmogorov en aquest cas es pot expressar

$$p(x;t,B)=\int_{\mathbb{R}} p(y;t-r,B)\cdot p(x;r,dy) \quad ,$$

o també

$$p(x;s+t,B)=\int_{\mathbb{R}} p(y;s,B)\cdot p(x;t,dy) \quad .$$

7.2.1 Exemples

1) Procés de Wiener.
   És un procés de Markov homogeni amb funció de transició

$$p(x;t,B)=\int_B \frac{1}{\sqrt{2\pi t}}\exp\{\frac{-(y-x)^2}{2t}\}\,dy \quad , \quad t\in\mathbb{R}^+,\ x\in\mathbb{R} \quad .$$

2) Procés d'Ornstein–Uhlenbeck.
   És un procés de Markov homogeni amb funció de transició

$$p(x;t,B)=\int_B \frac{1}{\sqrt{2\pi(1-e^{-2\gamma t})}}\exp\Big\{-\frac{(y-xe^{-\gamma t})^2}{2(1-e^{-2\gamma t})}\Big\} \quad .$$

3) Procés de Cauchy.
   El *procés de Cauchy* començant en 0 és el que compleix
   1) $X_0=0$
   2) $X_t-X_s\sim \text{Cauchy}(t-s)$ (densitat $\frac{t-s}{\pi((t-s)^2+x^2)}$).
   3) Té increments independents del passat.

   Es tracta d'un procés de Markov homogeni amb funció de transició

$$p(x;t,B)=\int_B \frac{t}{\pi(t^2+(x-y)^2)}\,dy \quad , \quad t\in\mathbb{R}^+,\ x\in\mathbb{R} \quad .$$

4) Cadenes de Markov.
   S'anomena *cadena de Markov* a tot procés de Markov que pren una quantitat numerable de valors. Podem suposar sense perdre generalitat que l'espai d'estats és $(\mathbb{N},\mathcal{P}(\mathbb{N}))$.
   Donar una funció de transició homogènia $p(\cdot;t,\cdot)\colon\mathbb{N}\times\mathcal{P}(\mathbb{N})\to\mathbb{R}$ és equivalent a donar les funcions

$$\begin{aligned} p_t\colon \mathbb{N}\times\mathbb{N} &\longrightarrow \mathbb{R}\\ (i,j) &\longmapsto p(i;t,\{j\}) \end{aligned}$$

   per $t\in\mathbb{R}^+$.
   L'equació de Chapman–Kolmogorov s'expressa

$$p_t(i,j)=\sum_{k\in\mathbb{N}} p_{t-r}(k,j)p_r(i,k) \quad .$$

   També es poden considerar cadenes de Markov no homogènies.

3) Procés de Poisson:
Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de variables aleatòries independents idènticament distribuïdes amb llei $\mathrm{Exp}(\lambda)$.
Sigui $S_n = \sum_{i=1}^{n} X_k$, $n \in \mathbb{N}$.
Definim $N_t := \sup\{n : \; S_n \leqslant t\}$, $t \in \mathbb{R}^+$.
Llavors, N_t té llei $\mathrm{Pois}(\lambda t)$:

$$P\{N_t = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad .$$

El procés $\{N_t, \; t \in \mathbb{R}+\}$ s'anomena *procés de Poisson amb intensitat* $\lambda > 0$, té increments estacionaris i independents del passat. És una cadena de Markov amb funció de transició homogènia donada per

$$p_t(i,j) = e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!} \quad , \quad j \geqslant i \quad .$$

7.3 Semigrup associat a funcions de transició markovianes homogènies

Posarem el contingut d'aquest apartat en termes més rigorosos més endavant. Introduïm només motivacions i exemples.
Sigui p una probabilitat de transició de l'espai mesurable $(E, \mathfrak{E})$ en ell mateix.
Sigui $b\mathfrak{E}$ l'espai vectorial de les funcions $E \to \mathbb{R}$ $\mathfrak{E}$-mesurables fitades.
p definexi una aplicació lineal

$$\begin{aligned} T\colon b\mathfrak{E} &\longrightarrow b\mathfrak{E} \\ f &\longmapsto [Tf](x) := \textstyle\int_E f(y)\, p(x,dy) \end{aligned}$$

Podem pensar que hem "estès" l'aplicació $p\colon E \times \mathfrak{E} \to \mathbb{R}$ a $p\colon E \times b\mathfrak{E} \to \mathbb{R}$ definida

$$p(x,f) := [Tf](x) = \int_E f(y)\, p(x,dy) \quad .$$

Identificant els indicadors amb els conjunts que indiquen, veiem que és certament una extensió.

Ara, si tenim una funció de transició markoviana $p(s,x;t,B)$ d'un procés a valors en $(E,\mathfrak{E})$, per cada $0 \leqslant s < t$ tindrem l'aplicació $T_{s,t}\colon b\mathfrak{E} \to b\mathfrak{E}$ definida per

$$[T_{s,t}f](x) = \int_E f(y)\, p(s,x;t,dy)$$

i també podem notar

$$p(s,x;t,f) := [T_{s,t}f](x) \quad .$$

Si la funció de transició és homogènia, aleshores la funció $p(x;t,B)$ defineix una aplicació $b\mathfrak{E} \to b\mathfrak{E}$ per

$$[T_t f](x) := \int_E f(y)\, p(x;t,dy)$$

i la família $\{T_t,\ t \in \mathbb{R}^+\}$ compleix la "propietat de semigrup"

$$T_s T_t f = T_{s+t} f \quad .$$

En efecte,

$$\begin{aligned} T_s T_t f = T_s(T_t f) &= T_s\Big(\int_E f(z)\, p(\cdot; t, dz)\Big) \\ &= \int_E \int_E f(z)\, p(y; t, dz)\, p(\cdot; s, dy) \\ &= \int_E f(z) \Big(\int_E p(y; t, dz)\, p(\cdot; s, dy)\Big) \\ &= \int_E f(z)\, p(\cdot; s+t, dz) = T_{s+t} f \quad . \end{aligned}$$

7.3.1 Exemple (Cadenes de Markov homogènies)
Aquí, $E = \mathbb{N}$ i $\mathfrak{E} = \mathcal{P}(\mathbb{N})$. Recordem que, en el cas de cadenes de Markov homogènies, podem donar com a funció de transició la família $p_t \colon \mathbb{N} \times \mathbb{N} \to \mathbb{R}$, $t \geqslant 0$. (Exemple 7.2.1). Tindrem que $b\mathfrak{E}$ és l'espai de les succesions reals, i que les aplicacions T_t són

$$[T_t f](i) = \sum_k f(k) \cdot p_t(i, k) \quad ,$$

que és el producte de la fila i-èsima de la matriu (possiblement infinita) p_t pel vector (possiblement infinit (successió)) f. O sigui, les aplicacions T_t es poden pensar com matrius (possiblement infinites).

7.3.2 Exemple (Cadenes de Markov amb espai d'estats finit)
Aquest exemple ens el podriem saltar. Formula pel cas particular d'espais d'estat finit el mateix que després trobarem en una situació molt més general. Però aquest cas és molt més fàcil d'entendre, evidentment.
A l'espai $\mathbb{R}^N$ ($N \in \mathbb{N}$), una aplicació lineal $T \colon \mathbb{R}^N \to \mathbb{R}^N$ ve representada unívocament per una matriu $N \times N$. Definim la norma d'una matriu com

$$\|A\| := \sup_{|x|=1} |Ax| \qquad (7.3.1) \quad ,$$

on $|\cdot|$ és la norma que fixem en $\mathbb{R}^N$. (Per exemple, si $|\cdot|$ és la norma euclidiana, la norma matricial associada és $\|A\| = \sqrt{\lambda_1}$, on λ_1 és el valor propi de mòdul més gran de la matriu simètrica $A^{\perp} A$.) El conjunt de les matrius $N \times N$ amb la suma, el producte per escalars i el producte de matrius, junt amb la norma (7.3.1), és una àlgebra de Banach.
La sèrie formal $\sum_{n \geqslant 0} \frac{x^n}{n!}$ és absolutament convergent en tot $\mathbb{R}$. Donada una aplicació lineal (matriu) T, la sèrie $\sum_{n \geqslant 0} \frac{T^n}{n!}$ és normalment convergent, doncs

$$\|T^n\| \leqslant \|T\| \|T^{n-1}\| \leqslant \cdots \leqslant \|T\|^n \quad .$$

Té sentit doncs definir l'operador (matriu)

$$\exp\{T\} := \sum_{n=0}^{\infty} \frac{T^n}{n!} \quad .$$

Aquest operador compleix:

1) $\exp\{T + S\} = \exp\{T\} \exp\{S\} \Leftrightarrow TS = ST$.
2) $\dfrac{d}{dt}\big(\exp\{tT\}\big) = T \exp\{tT\} = \exp\{tT\} T$.

Anem a veure una representació explícita de totes les possibles famílies $\{p_t,\ t \in \mathbb{R}^+\}$, suposant que es compleix $\lim\limits_{t\to 0} p_t = \mathrm{Id}$.

1r pas: La funció $t \mapsto p_t$ és contínua:

Sabem que $p_{t+h} = p_t p_h$ (propietat de semigrup). Llavors,

$$\lim_{h\to 0^+} p_{t+h} = p_t \lim_{h\to 0^+} p_h = p_t \quad ,$$

i

$$\lim_{h\to 0^+} p_{t-h} = \lim_{h\to 0+} p_t p_h^{-1} = p_t \lim_{h\to 0^+} p_h^{-1} = p_t \quad ,$$

tenint en compte que per h prou petit, p_h ha de ser invertible.

2n pas: La funció $t \mapsto p_t$ és derivable en zero:

Volem veure que existeix el límit

$$p'_{0^+} = \lim_{h\to 0+} \frac{p_h - \mathrm{Id}}{h} \quad . \tag{7.3.2}$$

Per a tot $h > 0$ i per a tot $n \in \mathbb{N}$, es té

$$\frac{p_h - \mathrm{Id}}{h} \sum_{k=0}^{n-1} p_{kh} h = p_{nh} - \mathrm{Id} \quad .$$

Usant la continuïtat de $t \mapsto p_t$, el segon factor convergeix, quan $h \to 0^+$ i $nh \to t$, a la integral de Riemann

$$\int_0^t p(s)\, ds \quad ,$$

que és invertible per t prou petit.

Per tant, el límit (7.3.2) existeix i val

$$p'_{0^+} = \big(p_t - \mathrm{Id}\big)\Big(\int_0^t p(s)\, ds\Big)^{-1} \quad ,$$

on t és qualsevol prou petit.

3r pas: La funció $t \mapsto p_t$ és derivable: Per $t > 0$,

$$p'_{t^+} = \lim_{h\to 0^+} \frac{p_{t+h} - p_t}{h} = \lim_{h\to 0^+} \frac{p_h - \mathrm{Id}}{h} \cdot p_t = p'_{0^+} p_t$$

Mateix resultat per l'esquerra.

4t pas: Forma explícita:

Notant A la matriu p'_{0^+}, hem obtingut que p_t satisfà l'equació diferencial

$$p'_t = A p_t$$

amb condició inicial $p_0 = \mathrm{Id}$. La única solució és

$$p_t = \exp tA \quad .$$

Observem que la matriu A compleix:

1) $A_{ij} \geqslant 0$ si $i \neq j$.
2) $\sum_{j=1}^{n} A_{ij} = 0$, per a tot $i = 1, \dots, n$.

La primera afirmació, perquè, per $i \neq j$,

$$A_{ij} = \lim_{h \to 0^+} \frac{p_h(i,j)}{h} \geqslant 0 \quad .$$

La segona, perquè $\sum_{j=1}^{n} p_t(i,j) = 1$.

Ens podem preguntar si, donada una matriu A complint les dues observacions anteriors, la funció $p_t = \exp\{tA\}$ és la funció de transició estacionària d'una cadena de Markov amb espai d'estats finit. La resposta és afirmativa:
Hem de veure que $p_t(i, \cdot)$ defineix una probabilitat en $\{1, \dots, N\}$ (o sigui, que $\sum_{j=1}^{n} p_t(i,j) = 1$ i que $p_t(i,j) \geqslant 0$).
Notem per $\mathbf{1}$ el vector amb 1 a totes les components.

$$p_t \cdot \mathbf{1} = \exp tA \cdot \mathbf{1} = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!} \cdot \mathbf{1} \quad . \tag{7.3.3}$$

Però $A \cdot 1 = 0$ i, iterativament, trobem $A^n \cdot \mathbf{1} = 0$, per a tot $n \geqslant 1$. Per tant, (7.3.3) és igual a $t^0 G^0 \mathbf{1} = \mathbf{1}$, i cada fila de p_t suma 1.
Per veure la positivitat de $p_t(i,j)$, escrivim el desenvolupament de Taylor

$$p_t(i,j) = \delta_{ij} + tA_{ij} + o(t) \quad .$$

Suposem primer que $A_{ij} > 0$. Està clar que per t prou petit, $p_t(i,j) > 0$. I com que

$$p_t = \exp\{tA\} = \exp\{\frac{t}{k} A\}^k = [p_{\frac{t}{k}}]^k \quad ,$$

això implica el mateix per a tot t.
Suposem que $A_{ij} = 0$. Sigui A^ε la mateixa matriu A però amb $A^\varepsilon_{ij} = \varepsilon > 0$. Aplicant l'argument anterior amb A^ε obtenim la positivitat de la $p_t(i,j)$ pertorbada, i fent $\varepsilon \to 0$ es veu que aquesta convergeix a la $p_t(i,j)$ sense pertorbar, que serà per tant positiva.
La matriu A s'anomena el *generador infinitesimal* de $\{p_t,\ t \in \mathbb{R}^+\}$.

7.4 Família markoviana homogènia respecte una filtració

Quan parlem de funcions de transició homogènies hem vist que podem estalviar-nos un argument, i considerar-les com a funció de tres variables. Quan parlem de famílies markovianes associades a una funció de transició homogènia hi ha un altre paràmetre superflu respecte la definició general: És suficient considerar la família de probabilitats $\{P_x\}_{x \in \mathbb{R}^d}$ en lloc de la família $\{P_{s,x}\}_{s \in \mathbb{R}^+, x \in \mathbb{R}^d}$.

Que es pugui fer això efectivament depèn de què poguem introduir els operadors shift (vegeu l'Apartat 6.3). Sabem que sempre és possible si estem a l'espai canònic.

Comprovem que ens podem estalviar el paràmetre s: Suposem que per a cada $\omega \in \Omega$ i cada $h > 0$ existeix $\omega' \in \Omega$ tal que $X_{t+h}(\omega) = X_t(\omega')$, $\forall t$. Aleshores, de la fórmula de la Proposició 6.1.3, deduirem que $\forall A \in \mathfrak{F}(X)$,

$$P_{s,x}(\theta_s^{-1} A) = P_{0,x}(A) \quad . \tag{7.4.1}$$

En efecte, ambdós membres són mesures com a funció de A i coincideixen sobre els conjunts de la forma $A = \{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\}$, $B_1, \dots, B_n \in \mathfrak{B}(\mathbb{R}^d)$, puix que

$$\begin{aligned}
&P_{s,x}(\theta_s^{-1}\{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\}) = P_{s,x}\{X_{t_1+s} \in B_1, \dots, X_{t_n+s} \in B_n\} \\
&= \int_{B_1}\int_{B_2}\cdots\int_{B_n} p(t_{n-1}+s, y_{n-1}; t_n+s, dy_n)\cdot\ldots\cdot p(s,x;t_1+s,dy_1) \\
&= \int_{B_1}\int_{B_2}\cdots\int_{B_n} p(t_{n-1}, y_{n-1}; t_n, dy_n)\cdot\ldots\cdot p(0,x;t_1,dy_1) \\
&= P_{0,x}\{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\}
\end{aligned}$$

Com que els conjunts d'aquesta forma formen un π-sistema que genera $\mathfrak{F}(X)$, la igualtat (7.4.1) és vàlida per a tot $A \in \mathfrak{F}(X)$.

Conclusió: Les probabilitats $P_{s,x}$ queden determinades per les probabilitats $P_{0,x}$.

Ara copiem més o menys la Definició 6.1.1.

7.4.1 Definició

Sigui Ω un conjunt (de moment, no hi posem cap σ-àlgebra ni cap probabilitat).
Sigui, per a cada $t \in \mathbb{R}^+$, una aplicació $X_t \colon \Omega \to \mathbb{R}^d$.
Suposem que per a cada $\omega \in \Omega$ i cada $h > 0$, existeix $\omega' \in \Omega$ tal que $X_{t+h}(\omega) = X_t(\omega')$, $\forall t$.
Sigui $p(x; t, B)$ una funció de transició homogènia:

1) $p(\cdot; t, B)$ és mesurable, $\forall t \in \mathbb{R}^+$, $\forall B \in \mathfrak{B}(\mathbb{R}^d)$.
2) $p(x; t, \cdot)$ és una probabilitat en $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$, $\forall t$, $\forall x \in \mathbb{R}^d$.
3) $p(x; 0, \cdot) = \delta_x$, $\forall s \in \mathbb{R}^+$, $\forall x \in \mathbb{R}^d$.

Per a tot $x \in \mathbb{R}^d$, sigui P_x una probabilitat sobre la σ-àlgebra de Ω generada per $\{X_t,\ t \geqslant 0\}$, que denotarem $\mathfrak{F}(X)$.
Diem que $\{X_t,\ t \in \mathbb{R}^+\}$, $\{P_x,\ x \in \mathbb{R}^d\}$ és una *família markoviana homogènia* amb funció de transició p si

1) $\forall x \in \mathbb{R}^d$, la família d'aplicacions $\{X_t,\ t \geqslant 0\}$ és un procés de Markov homogeni a l'espai de probabilitat $(\Omega, \mathfrak{F}(X), P_x)$ amb funció de transició p.
2) $P_x\{X_0 = x\} = 1$.

7.4.2 Teorema

Sigui p una funció de transició homogènia complint l'equació de Chapman–Kolmogorov

$$p(x; s+t, B) = \int_{\mathbb{R}^d} p(y; s, B)\, p(x; t, dy) \quad .$$

Aleshores existeix una família markoviana homogènia amb funció de transició p.

Demostració: Com la del Teorema 6.1.6. La construcció es fa sobre l'espai canònic, de manera que la condició d'existència de ω' a partir de ω se satisfà trivialment.

7.5 Expressions de la propietat de Markov amb els operadors shift

Tenim per al cas homogeni un anàleg de la fórmula

$$P_{u,x}\big(A/_{\mathfrak{F}_s(X)}\big) = P_{s,X_s}(A) \quad , \quad P_{X_s}\text{-q.s.}, \quad \forall A \in \mathfrak{F}^s(X)$$

(Corol·lari 6.1.4):

Com que, segons hem vist a l'apartat anterior,

$$P_{s,x}(\theta_s^{-1}A) = P_{0,x}(A) \quad , \quad \forall A \in \mathfrak{F}(X) \quad ,$$

l'anàleg serà

$$P_x\big(\theta_s^{-1}A/_{\mathfrak{F}_s(X)}\big) = P_{X_s}(A) \quad , \quad P_{X_s} - \text{q.s.}, \quad \forall A \in \mathfrak{F}^s(X) \quad ,$$

que té una interpretació molt fàcil: El procés a l'instant t, suposant conegut el seu passat fins s, es comporta com si comencés de l'instant 0 partint del punt $X_s(\omega)$.

8. Instants d'aturada

8.1 Instants d'aturada i instants opcionals

Sigui $(\Omega, \mathfrak{F})$ un espai mesurable.

Sigui $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ una filtració en aquest espai (és a dir, una família creixent de sub-σ-àlgebres de $\mathfrak{F}$.

Es defineixen les σ-àlgebres

$$\mathfrak{F}_{t^+} := \bigcap_{s>t} \mathfrak{F}_s \quad , \quad t \geqslant 0 \quad ,$$

$$\mathfrak{F}_{t^-} := \bigvee_{s<t} \mathfrak{F}_s \quad , \quad t > 0 \quad .$$

Diem que la filtració $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ és *contínua per la dreta* (càd) si $\forall t$, $\mathfrak{F}_t = \mathfrak{F}_{t^+}$.

Observeu que $\{\mathfrak{F}_{t^+},\ t \in \mathbb{R}^+\}$ és també una filtració, i que és contínua per la dreta.

8.1.1 Definició

Sigui $(\Omega, \mathfrak{F})$ un espai mesurable.
Un *instant aleatori* és una funció mesurable $\tau \colon \Omega \to [0, +\infty]$.

8.1.2 Definició

Sigui $(\Omega, \mathfrak{F})$ un espai mesurable.
Sigui $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ una filtració.
Un instant aleatori τ tal que

$$\forall t \in \mathbb{R}^+, \quad \{\tau \leqslant t\} \in \mathfrak{F}_t \quad ,$$

s'anomena *instant d'aturada* respecte la filtració.
En el context dels processos de Markov, de vegades s'anomenen *instants markovians*.

8.1.3 Definició

En la situació anterior, si

$$\forall t \in \mathbb{R}^+, \quad \{\tau < t\} \in \mathfrak{F}_t \quad ,$$

aleshores τ és un *instant opcional* (o *instant d'aturada en sentit ampli*) respecte la filtració.

8.1.4 Observacions

1) Si τ és constant, aleshores és instant d'aturada respecte qualsevol filtració.

2) Suposem que $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ i $\{\mathfrak{G}_t,\ t \in \mathbb{R}^+\}$ són filtracions tals que $\mathfrak{F}_t \subset \mathfrak{G}_t$, $\forall t \in \mathbb{R}^+$. Si τ és instant d'aturada (resp. opcional) respecte $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$, aleshores ho és també respecte $\{\mathfrak{G}_t,\ t \in \mathbb{R}^+\}$.
3) Si τ pren un quantitat numerable de valors, és instant d'aturada si i només si
$$\forall t \in \mathbb{R}^+, \quad \{\tau = t\} \in \mathfrak{F}_t \quad .$$
4) Si τ és instant d'aturada, aleshores és instant opcional.
5) τ és instant opcional si i només si és instant d'aturada respecte $\{\mathfrak{F}_{t^+},\ t \in \mathbb{R}^+\}$.
6) Si la filtració és càd, instant d'aturada i instant opcional són equivalents.
7) Si τ és instant opcional i a és una constant estrictament positiva, aleshores $\tau + a$ és instant d'aturada. $\square$

Totes les observacions anteriors són fàcils de demostrar.

8.1.5 Exercici

Un instant opcional no té perquè ser instant d'aturada. Es pot construir un exemple mitjançant un procés amb dues trajectòries contínues.

8.1.6 Proposició

Si τ i ρ són instants d'aturada (resp. opcionals), aleshores també ho són $\tau \wedge \rho$ (ínfim de τ i ρ), $\tau \vee \rho$ (suprem de τ i ρ), i $\tau + \rho$.

Demostració:

a) $\{\tau \vee \rho \leqslant t\} = \{\tau \leqslant t\} \cap \{\rho \leqslant t\} \in \mathfrak{F}_t$.
b) $\{\tau \wedge \rho > t\} = \{\tau > t\} \cap \{\rho > t\} \in \mathfrak{F}_t$.
c)
$$\{\tau + \rho > t\} =$$
$$\{\tau = 0,\ \rho > t\} \cup \{\tau > t,\ \rho = 0\} \cup \{\tau \geqslant t,\ \rho > 0\} \cup \{0 < \tau < t,\ \tau + \rho > t\}$$

Els tres primers conjunts són interseccions de conjunts de $\mathfrak{F}_t$. El quart és igual a la unió, per $r \in \mathbb{Q}$ i $0 < r < t$, de la família numerable de conjuunts de $\mathfrak{F}_t$
$$\{r < \tau < t,\ \rho > t - r\} \quad . \quad \square$$

8.1.7 Proposició

Si $\{\tau_n\}_{n\in\mathbb{N}}$ és una successió d'instants opcionals, aleshores també ho són
$$\sup_{n\in\mathbb{N}} \tau_n \quad , \quad \inf_{n\in\mathbb{N}} \tau_n \quad , \quad \overline{\lim_{n\in\mathbb{N}}}\, \tau_n \quad , \quad \underline{\lim_{n\in\mathbb{N}}}\, \tau_n \quad .$$

Demostració: Per l'ínfim, observeu que
$$\Big\{ \inf_{n\in\mathbb{N}} \tau_n < t \Big\} = \bigcup_{n=1}^{\infty} \{\tau_n < t\} \in \mathfrak{F}_t \quad .$$

Pel suprem,
$$\Big\{ \sup_{n\in\mathbb{N}} \tau_n \leqslant t \Big\} = \bigcap_{n=1}^{\infty} \{\tau_n \leqslant t\} \in \mathfrak{F}_{t^+} \quad ,$$

d'on $\sup_{n\in\mathbb{N}} \tau_n$ és un instant d'aturada respecte $\{\mathfrak{F}_{t^+},\ t \in \mathbb{R}^+\}$, i per tant és opcional respecte $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$.
De passada, observem que la mateixa igualtat ens dóna que si els τ_n són instants d'aturada, aleshores $\sup_{n\in\mathbb{N}} \tau_n$ és instant d'aturada. Amb els ínfims això no funciona.
El resultat per límits superiors i inferiors són immediats a partir d'aquests. □

8.1.8 Proposició

Sigui τ un instant d'aturada.
Aleshores existeix una successió d'instants d'aturada $\{\tau_n\}_{n\in\mathbb{N}}$ tal que cada τ_n pren una quantitat finita de valors i $\tau_n \searrow \tau$.

Demostració: Funciona la successió següent:

$$\tau_n(\omega) = \begin{cases} 0\ , & \text{si } \tau(\omega) = 0\ . \\ \frac{k}{2^n}\ , & \text{si } \frac{k-1}{2^n} < \tau(\omega) \leqslant \frac{k}{2^n}\ . \\ \infty\ , & \text{si } \tau(\omega) > n\ . \quad \square \end{cases}$$

La idea intuïtiva rera els conceptes de filtració i d'instant d'aturada és la següent:
Una filtració $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$ s'interpreta com una acumulació d'informació que va augmentant a mesura que el temps flueix. La informació de què es disposa a l'instant t és la σ-àlgebra $\mathfrak{F}_t$. Això vol dir que a l'instant t podem decidir si s'ha produït o no qualsevol esdeveniment contingut en $\mathfrak{F}_t$, i cap altre. Per exemple, si l'acumulació d'informació és deguda exclusivament a la observació d'un procés estocàstic $\{X_t,\ t \in \mathbb{R}^+\}$, la filtració a considerar és la "filtració natural" del procés: $\mathfrak{F}_t = \sigma\{X_s,\ s \leqslant t\}$. Està clar que a l'instant t només podem preguntar-nos per la realització d'esdeveniments que involucrin (i.e. de la σ-àlgebra generada per) les variables aleatòries $\{X_s,\ s \leqslant t\}$.
Suposem que volem parlar del moment en què s'ha produït un cert fenomen, i que aquest moment depèn de $\omega \in \Omega$; és a dir, és un instant aleatori. L'exemple típic (i més important) és el moment en què un procés estocàstic entra per primera vegada en un cert conjunt B de l'espai d'estats. Podem expressar-lo com

$$\tau(\omega) = \min\{s :\ X_s(\omega) \in B\}$$

(suposem de moment que per a cada ω aquest mínim existeix realment). És força intuïtiu que l'esdeveniment $\{\omega \in \Omega :\ \tau(\omega) \leqslant t\}$, que es realitza si i només si el procés ha entrat en B abans de o en l'instant t, podem decidir si s'ha realitzat o no quan estem a l'instant t. En altres paraules, ha de ser part de la informació acumulada fins a l'instant t, o sigui, ha de pertànyer a $\mathfrak{F}_t$. I això és el que demanem a la definició.

Una altra imatge molt descriptiva per entendre els instants d'aturada és: Suposem que estem a l'espai canònic, $\Omega = \mathbb{R}^{\mathbb{R}^+}$. Una variable aleatoria $\tau\colon \Omega \to [0,\infty]$ és un instant d'aturada si $\forall \omega_1, \omega_2 \in \Omega$ tals que $\omega_1(s) = \omega_2(s)$, $\forall 0 \leqslant s \leqslant t$, $\tau(\omega_1) \leqslant t$ implica $\tau(\omega_1) = \tau(\omega_2)$. Heurística: L'ocorrència o no de $\{\omega :\ \tau(\omega) \leqslant t\}$ es pot establir a partir de l'observació del procés en $[0,t]$ només.

8.2 Esdeveniments anteriors a un instant d'aturada

Ara podem preguntar-nos si podem formalitzar la informació acumulada fins a un instant d'aturada, semblantment a com $\mathfrak{F}_t$ formalitza la informació acumulada fins a l'instant determinista t.

Suposem que A és un esdeveniment que forma part d'aquesta informació fins a l'instant d'aturada τ. Això vol dir: A l'instant τ, podem decidir si A s'ha realitzat o no.

Suposem que a l'instant t coneixem el valor de τ (cosa que només és possible si $\tau \leqslant t$). Aleshores, hem de ser capaços també de saber si A s'ha realitzat o no. Estem dient, doncs, que els esdeveniments

$$A \cap \{\tau \leqslant t\} \qquad \text{i} \qquad A^c \cap \{\tau \leqslant t\}$$

han de ser de la σ-àlgebra $\mathfrak{F}_t$. A més si això passa per un dels dos, també passa per l'altre, perquè

$$A^c \cap \{\tau \leqslant t\} = \big(A \cap \{\tau \leqslant t\}\big)^c \cap \{\tau \leqslant t\} \quad .$$

Tot això ens porta a la definició següent.

8.2.1 Definició

Sigui τ un instant d'aturada relatiu a $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$.
S'anomena *σ-àlgebra dels esdeveniments anteriors a* τ la formada pels $A \in \mathfrak{F}$ tals que

$$A \cap \{\tau \leqslant t\} \in \mathfrak{F}_t \quad , \quad \forall t \geqslant 0 \quad .$$

S'escriu $\mathfrak{F}_\tau$.

8.2.2 Observacions

1) τ és $\mathfrak{F}_\tau$-mesurable.
2) Si $\tau \equiv t$, llavors $\mathfrak{F}_\tau = \mathfrak{F}_t$.

8.2.3 Proposició

Siguin τ, ρ instants d'aturada.
Aleshores:

1) $A \in \mathfrak{F}_\rho \Rightarrow A \cap \{\rho \leqslant \tau\} \in \mathfrak{F}_\tau$ *(en particular,* $\tau \leqslant \rho \Rightarrow \mathfrak{F}_\rho \subset \mathfrak{F}_\tau$*).*

2) $\mathfrak{F}_{\tau\wedge\rho} = \mathfrak{F}_\tau \cap \mathfrak{F}_\rho$ *i els esdeveniments*

$$\{\tau \leqslant \rho\} \quad , \quad \{\tau < \rho\}$$

hi pertanyen.

Demostració:

1) Sigui $A \in \mathfrak{F}_\rho$. Llavors

$$A \cap \{\rho \leqslant \tau\} \cap \{\tau \leqslant t\} = [A \cap \{\rho \leqslant t\}] \cap \{\tau \leqslant t\} \cap \{\rho \wedge t \leqslant \tau \wedge t\} \quad .$$

Els tres conjunts són de $\mathfrak{F}_t$, el tercer d'ells perquè per a qualssevol instant d'aturada τ i instant determinista t, la variable aleatòria $\tau \wedge t$ és $\mathfrak{F}_t$-mesurable (fàcil de comprovar).

2) La inclusió $\mathfrak{F}_{\tau\wedge\rho} \subset \mathfrak{F}_\tau \cap \mathfrak{F}_\rho$ surt de $\tau \wedge \rho \leqslant \tau$, $\tau \wedge \rho \leqslant \rho$ i la part 1).
Per a la inclusió contrària, si $A \in \mathfrak{F}_\tau \cap \mathfrak{F}_\rho$, llavors

$$A \cap \{\tau \wedge \rho \leqslant t\} = A \cap \big[\{\tau \leqslant t\} \cup \{\rho \leqslant t\}\big] \quad ,$$

que és de $\mathfrak{F}_t$.
De 1) aplicat a $A = \Omega$ obtenim que $\{\tau \leqslant \rho\} \in \mathfrak{F}_\rho$. L'instant d'aturada $\tau \wedge \rho$ és, també per 1), $\mathfrak{F}_\rho$-mesurable. Però $\{\tau < \rho\} = \{\tau \wedge \rho < \rho\} \in \mathfrak{F}_\rho$.

Sigui $\{X_t,\ t \in \mathbb{R}^+\}$ un procés estocàstic i τ un instant aleatori. Aleshores podem definir sobre $\{\tau < \infty\}$ la funció

$$X_\tau(\omega) := X_{\tau(\omega)}(\omega) \quad .$$

8.2.4 Exercici

Si $\{X_t,\ t \in \mathbb{R}^+\}$ és un procés estocàstic mesurable i τ és finit, aleshores X_τ és una variable aleatòria en $(\Omega, \mathfrak{F}, P)$.

8.3 Procés aturat

8.3.1 Definició

Sigui $\{X_t,\ t \in \mathbb{R}^+\}$ un procés estocàstic.
Sigui τ un instant aleatori.
Definim el *procés X aturat per τ* com $\{X_{t\wedge\tau},\ t \in \mathbb{R}^+\}$.

8.3.2 Definició

Un procés estocàstic $\{X_t,\ t \in \mathbb{R}^+\}$ s'anomena *progressivament mesurable* (o *progressiu*) si per a tot $t \in \mathbb{R}^+$, la restricció

$$X_{|_{[0,t]\times\Omega}} \colon [0,t] \times \Omega \longrightarrow \mathbb{R}$$

és mesurable respecte la σ-àlgebra producte $\mathfrak{B}([0,t]) \otimes \mathfrak{F}_t$. $\square$

Es veu fàcilment que tot procés progressivament mesurable és mesurable. El recíproc no és cert.

8.3.3 Proposició

Sigui $\{X_t,\ t \in \mathbb{R}^+\}$ progressivament mesurable.
Sigui τ un instant d'aturada.
Aleshores:

1) Si τ és finit, la variable aleatòria X_τ és $\mathfrak{F}_\tau$-mesurable.
2) El procés X aturat per τ és progressivament mesurable.

Demostració:

1) Volem veure que, per a tot $B \in \mathfrak{B}(\mathbb{R})$, $\{X_\tau \in B\} \cap \{\tau \leqslant t\} \in \mathfrak{F}_t$. Aquesta intersecció es pot escriure com

$$\big\{\omega \in \Omega :\ (\tau(\omega), \omega) \in \{(s,\omega) :\ s \leqslant t,\ X_s(\omega) \in B\}\big\}$$

El conjunt

$$\{(s,\omega) :\ s \leqslant t,\ X_s(\omega) \in B\}$$

és de $\mathfrak{B}([0,t]) \otimes \mathfrak{F}_t$, per la progressivitat del procés.
Haurem acabat si veiem que

$$\forall A \in \mathfrak{B}([0,t]) \otimes \mathfrak{F}_t, \quad \{\omega \in \Omega :\ (\tau(\omega),\omega) \in A\} \in \mathfrak{F}_t \quad .$$

La col·lecció de conjunts A pels quals això és cert conté els conjunts del tipus $[0,s] \times B$, que generen la σ-àlgebra $\mathfrak{B}([0,t]) \otimes \mathfrak{F}_t$.

2) Només cal escriure l'aplicació

$$(s,\omega) \longmapsto X_{s\wedge\tau}(\omega)$$

com a composició de

$$(s,\omega) \longmapsto (s\wedge\tau,\omega)$$

i

$$(s,\omega) \longmapsto X_s(\omega) \quad .$$

La primera és $\big(\mathfrak{B}([0,t])\otimes\mathfrak{F}_t, \mathfrak{B}([0,t])\otimes\mathfrak{F}_t\big)$-mesurable, i la segona és $\big(\mathfrak{B}([0,t])\otimes\mathfrak{F}_t, \mathfrak{B}(\mathbb{R})\big)$-mesurable.

8.3.4 Exercici

La Propietat de la Torre de les esperances condicionades s'estén a instants d'aturada, de la manera següent: Si X és una variable aleatòria integrable, i τ i ρ són instants d'aturada,

$$\mathrm{E}\left[\mathrm{E}\left[{}^{X}\!/_{\mathfrak{F}_\tau}\right]/_{\mathfrak{F}_\rho}\right] = \mathrm{E}\left[{}^{X}\!/_{\mathfrak{F}_{\tau\wedge\rho}}\right] \quad . \quad \square$$

Gairebé qualsevol llibre de processos estocàstics parla d'instants d'aturada (stopping times, temps d'arrêt) i d'instants opcionals. Vegeu per exemple en Karatzas–Shreve [5] algunes propietats més.

8.4 Entrades a conjunts

L'exemple més interessant de instants aleatoris és el que s'anomena *instant d'entrada d'un procés a un conjunt* $B\in\mathfrak{B}(\mathbb{R})$, que és per definició

$$\tau(\omega) = \inf\{t:\ X_t(\omega)\in B\} \quad .$$

Un instant d'entrada és instant d'aturada en certes situacions. Citem-ne tres:

1) Si $\{X_t,\ t\in\mathbb{R}^+\}$ és continu per la dreta i B és obert.
2) Si $\{X_t,\ t\in\mathbb{R}^+\}$ és continu i B és tancat.
3) Si $\{X_t,\ t\in\mathbb{R}^+\}$ és progressivament mesurable, $\{\mathfrak{F}_t,\ t\in\mathbb{R}^+\}$ és càd, i $(\Omega,\mathfrak{F}_0,P)$ és un espai de probabilitat complet.

Les dues primeres afirmacions valen per processos amb valors a qualsevol espai mètric. La tercera val per processos a valors en un espai polonès.

8.5 Aclariment sobre mesurabilitat, progressivitat i adaptació

Les relacions entre els conceptes de mesurabilitat, progressivitat, adaptació, i la continuïtat a la dreta (càd) o a l'esquerra (càg) de les trajectòries són les següents i només les següents:

1) càd o càg $\Rightarrow$ mesurable
2) progressiu $\Rightarrow$ mesurable
3) progressiu $\Rightarrow$ adaptat
4) adaptat i (càd o càg) $\Rightarrow$ progressiu

9. Família fortament Markov

Idea: Essencialment, els processos fortament markovians seran aquells que compleixen la propietat de Markov

$$P\left(X_{t+h} \in B/_{\mathfrak{F}_t}\right) = P\left(X_{t+h} \in B/_{X_t}\right)$$

no només pels instants deterministes s sino també per certs instants aleatoris.

Aquests instants aleatoris no poden ser qualssevol. Sempre en podem inventar que no farien independents passat i futur, per qualsevol procés de Markov (excepte trivials).

Hi ha presentacions molt diverses de la propietat forta de Markov. Segueixo Wentzell [12].

Fem una definició preliminar.

9.0.1 Definició

Si X és un procés estocàstic mesurable i τ és un instant aleatori, aleshores els conjunts de la forma

$$\{\omega \in \Omega : \ \tau(\omega) < \infty \text{ i } X_{\tau(\omega)}(\omega) \in B\}$$

amb $B \in \mathfrak{B}(\mathbb{R})$, juntament amb el conjunt $\{\tau = \infty\}$ constitueixen una σ-àlgebra, que s'anomena la *σ-àlgebra generada per X_τ*. $\square$

En el que segueix escrivim conjunts tals com $\{X_\tau \in B\}$ amb el significat $\{\tau < \infty, \ X_\tau \in B\}$

9.0.2 Proposició

Sigui $\{X_t, \ t \in \mathbb{R}^+\}$, $\{P_x\}_{x\in\mathbb{R}^d}$ una família markoviana homogènia, amb funció de transició p.
Sigui $\{\mathfrak{F}_t, \ t \geqslant 0\}$ una filtració.
Suposem que $\{X_t, \ \in \mathbb{R}^+\}$ és progressivament mesurable respecte la filtració $\{\mathfrak{F}_t, \ t \in \mathbb{R}^+\}$.
Aleshores, la funció de transició $p(x; t, B)$ és mesurable en $(x; t)$ respecte $\mathfrak{B}(\mathbb{R}^+) \otimes \mathfrak{B}(\mathbb{R}^d)$, $\forall B \in \mathfrak{B}(\mathbb{R}^d)$.

9.0.3 Definició

Una família markoviana homogènia $\{X_t, \ t \in \mathbb{R}^+\}$, $\{P_x\}_{x\in\mathbb{R}^d}$ és *fortament markoviana* respecte la filtració $\{\mathfrak{F}_t, \ t \in \mathbb{R}^+\}$ si

1) És progressivament mesurable respecte $\{\mathfrak{F}_t\}_{t\in\mathbb{R}^+}$.

2) Per a tot instant d'aturada τ, i per a tota funció

$$\eta\colon \{\tau < \infty\} \longrightarrow [0, \infty] \quad ,$$

mesurable respecte $\mathfrak{F}_\tau$, es té

$$P\bigl(X_{\tau+\eta} \in B /_{\mathfrak{F}_\tau}\bigr) = p(X_\tau; \eta, B) \quad , \quad P_x\text{-q.s.}, \ \forall x \in \mathbb{R}^d, \ \forall B \in \mathfrak{B}(\mathbb{R}^d) \quad \square \tag{9.0.1}$$

Comprovem que la definició té sentit: Per definició, (9.0.1) vol dir

$$P_x(A \cap \{X_{\tau+\eta} \in B\}) = \int_A p(X_\tau; \eta, B)\, P_x(d\omega) \quad ,$$

$\forall A \subset \{\tau, \eta < \infty\}$, $A \in \mathfrak{F}_\tau$.

La integral té sentit:

a) τ i η són mesurables respecte $\mathfrak{F}_\tau$.

b) $\{X_t,\ t \in \mathbb{R}^+\}$ és proegressivament mesurable, i per tant X_τ és mesurable respecte $\mathfrak{F}_\tau$ sobre $\{\tau < \infty\}$.

a) i b) impliquen que la funció $\omega \mapsto (X_\tau(\omega), \eta(\omega)$ és mesurable respecte $\mathfrak{F}_\tau$.

Per la proposició anterior, $(x; t) \mapsto p(x; t, B)$ és mesurable. Aplicant que la composició de mesurables és mesurable, obtenim la mesurabilitat de $p(X_\tau; \eta, B)$.

El que resulta de la definició, si ho apliquem a τ i η no aleatoris, és una família markoviana respecte $\{\mathfrak{F}_t,\ t \in \mathbb{R}^+\}$. Això no hem definit. Si es vol, podem suposar que la filtració és la natural del procés i ja estem en la situació coneguda.

Tota família fortament markoviana és doncs markoviana. El recíproc no és cert. Però es pot demostrar que la igualtat (9.0.1) és compleix sempre per a tota família markoviana si τ i η prenen una quantitat numerable de valors.

10. Trajectòries dels processos de Markov

Està clar que la definició de procés de Markov no implica absolutament res sobre la regularitat de les trajectòries. Per exemple, un procés amb una sola trajectòria és Markov, i aquesta trajectòria pot ser tan irregular com es vulgui.

El que volem són criteris que, a la vista de la funció de transició, ens donin propietats de regularitat de les trajectòries. De tota manera, el conjunt de trajectòries d'un procés de Markov no està únicament determinat per la funció de transició. La pregunta correcta és del tipus: "Existeix algun procés de Markov amb la funció de transició donada p tal que totes les trajectòries siguin contínues?". Equivalentment: "El conjunt de trajectòries contínues del procés canònic associat a la funció p donada, té mesura exterior 1?"

Comencem introduint l'espai de funcions contínues en $\mathbb{R}^+$, que es pot considerar "l'espai canònic" per als processos amb trajectòries contínues. Aquesta discussió lliga directament amb la de l'Apartat 1.4.

10.1 L'espai de funcions contínues en $\mathbb{R}^+$

Sigui $C = C(\mathbb{R}^+; \mathbb{R}^d)$ l'espai vectorial de les funcions contínues sobre $\mathbb{R}^+$, a valors en $\mathbb{R}^d$.

Sigui, per a cada $\omega \in C$, $X_t(\omega) = \omega(t)$.

Definim en C la topologia de la convergència uniforme sobre acotats (vegeu el Capítol 11). Aquesta topologia és metritzable mitjançant la distància

$$d(\omega, \omega') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{0 \leqslant t \leqslant n} |X_t(\omega) - X_t(\omega')|}{1 + \sup_{0 \leqslant t \leqslant n} |X_t(\omega) - X_t(\omega')|} \quad .$$

L'espai resultant és polonès.

Sigui $\mathfrak{B}(C)$ la σ-àlgebra de Borel d'aquest espai mètric.

10.1.1 Proposició

$$\sigma\{X_t,\ t \geqslant 0\} = \mathfrak{B}(C) \quad .$$

Demostració: L'aplicació $\omega \mapsto X_t(\omega)$ és contínua per a cada t (és el funcional lineal δ_t) i per tant mesurable respecte $\mathfrak{B}(C)$. Tenim doncs que $\sigma\{X_t,\ t \geqslant 0\} \subset \mathfrak{B}(C)$.

Per a la inclusió recíproca: Un sistema fonamental d'entorns de la topologia de C és el dels conjunts de la forma

$$\Big\{\omega:\ \sup_{0\leqslant t\leqslant T}|X_t(\omega)-X_t(\omega^0)|<\varepsilon\Big\} \tag{10.1.1}$$

variant ω^0, T, ε (les boles en cada $C([0,T];\mathbb{R}^d)$).

Per a cada ω^0, T, ε fixats, podem posar el conjunt (10.1.1) com

$$\bigcup_{n=1}^{\infty}\{\omega:\ |X_t(\omega)-X_t(\omega^0)|\leqslant\varepsilon(1-\frac{1}{n}),\ \forall t\in[0,T]\cap\mathbb{Q}\}\quad .$$

Cada conjunt d'aquesta unió numerable és de $\sigma\{X_t,\ t\geqslant 0\}$, de on (10.1.1) també, i la σ-àlgebra generada per aquests (10.1.1) estarà inclosa en $\sigma\{X_t,\ t\geqslant 0\}$. O sigui $\mathfrak{B}(C)=\sigma\{X_t,\ t\geqslant 0\}$. □

10.1.2 Exercici

De la continuïtat de les trajectòries es dedueix també que

$$\sigma\{X_r,\ r\geqslant t\}=\bigvee_{s<t}\sigma\{X_r,\ r\geqslant s\}\quad,\quad\forall t>0\quad .$$ □

Recordem que a l'Apartat 1.4 vèiem un procediment general per estendre l'espai canònic de manera que fos mesurable un conjunt Γ al qual pertanyen les trajectòries d'un procés amb probabilitat 1. Notarem $\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma]$ la σ-àlgebra estesa resultant.

Un cop tenim Γ mesurable i de probabilitat 1 en aquest nou espai $((\mathbb{R}^d)^{\mathbb{R}^+},\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma],P)$, podem deixar totalment de banda el conjunt Γ^c i restringir l'espai: S'agafa, en lloc de $(\mathbb{R}^d)^{\mathbb{R}^+}$, el conjunt Γ; en lloc de $\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma]$, la *traça* d'aquesta σ-algebra amb Γ (intersecar cada conjunt de $\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma]$ amb Γ; el que resulta és una σ-àlgebra que podem seguir notant per $\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma]$ sense confusió possible), i definir sobre aquesta traça

$$P^\Gamma(A\cap\Gamma)=P(A)\quad,\quad A\in\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma]\quad . \tag{10.1.2}$$

10.1.3 Proposició

En la situació anterior, P^Γ està ben definida i és la única mesura sobre $(\Gamma,\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma])$ complint (10.1.2).

Demostració:

Unicitat: Sigui Q complint (10.1.2). Q i P^Γ coincideixen sobre tot mesurable contingut en Γ. O sigui, $Q=P^\Gamma$ en $(\Gamma,\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma])$.

Existència: Si $A,B\in\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[\Gamma]$ compleixen $A\cap\Gamma=B\cap\Gamma$, aleshores $(A-B)\cup(B-A)\subset(\mathbb{R}^d)^{\mathbb{R}^+}-\Gamma$, i com Γ té mesura exterior 1, necessàriament $P\big((A-B)\cup(B-A)\big)=0$, el que implica $P(A)=P(B)$ i per tant P^Γ està ben definida. □

Considerem ara el cas en què $\Gamma=C$. Hi ha un resultat fàcil però molt important, que és el següent. De fet, és una reformulació de la Proposició 10.1.1.

10.1.4 Teorema

L'espai mesurable $(C,\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[C])$ és $(C,\mathfrak{B}(C))$.

Demostració: Només cal observar que $\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}[C]$ és la mínima σ-àlgebra en $\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}$ que fa mesurables les projeccions X_t, i el conjunt C. Per tant, fent la traça per C, resulta exactament $\sigma\{X_t,\ t\geqslant 0\}=\mathfrak{B}(C)$, segons la Proposició 10.1.1.

10.2 Criteris de continuïtat de les trajectòries

Hi ha un criteri general de continuïtat de les trajectòries d'un procés de Markov basat en el criteri de continuïtat de Kolmogorov per a processos qualssevol, que escrivim prèviament.

10.2.1 Teorema

Sigui P una probabilitat en $(\mathbb{R}^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R})^{\mathbb{R}^+})$.
Suposem que per a cada $T > 0$, existeixen $\alpha > 0$, $r > 0$ i $K > 0$ tals que

$$\mathrm{E}[|\omega(t) - \omega(s)|^r] \leqslant K|t-s|^{1+\alpha} \quad , \quad \forall 0 \leqslant s < t \leqslant T \quad . \tag{10.2.1}$$

Aleshores C té mesura exterior 1. □

D'aquí es dedueix:

10.2.2 Teorema

Sigui $\{P_{t_1,\dots,t_n},\ t_i \in \mathbb{R}^+,\ n \in \mathbb{N}\}$ una família consistent de lleis en dimensió finita.
Suposem que per a cada $T > 0$ existeixen $\alpha > 0$, $r > 0$, $K > 0$ tals que

$$\int_{\mathbb{R}^2} |y-x|^r \, P_{s,t}(d(x,y)) \leqslant K|t-s|^{1+\alpha} \quad , \quad \forall 0 \leqslant s < t \leqslant T \quad . \tag{10.2.2}$$

Aleshores existeix una única probabilitat P sobre $(C, \mathfrak{B}(C))$ tal que

$$P\{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\} = P_{t_1,\dots,t_n}(B_1 \times \cdots \times B_n) \quad ,$$

$\forall 0 \leqslant t_1 < \cdots < t_n,\ \forall B_1, \dots, B_n \in \mathfrak{B}(\mathbb{R}^d) \quad .$

Demostració: La unicitat ve del fet que $\mathfrak{B}(C) = \sigma\{X_t,\ t \geqslant 0\}$, on X_t és el procés coordenat, com abans.

L'existència: Existeix P' en $((\mathbb{R}^d)^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+})$ única determinada per la família $P_{t_1,\dots,t_n}$. Llavors, (10.2.2) implica (10.2.1) i C té mesura exterior 1 en $((\mathbb{R}^d)^{\mathbb{R}^+}, \mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}, P')$.

P' determina una probabilitat P al restringir-la a l'espai C, dotat amb la σ-àlgebra traça de $\mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}$ amb C, i definida per $P(A \cap C) = P'(A)$, $\forall A \in \mathfrak{B}(\mathbb{R}^d)^{\mathbb{R}^+}$. Però aquesta traça és $\mathfrak{B}(C)$, segons el Teorema 10.1.4.

10.2.3 Observació

Sota la hipòtesi del Teorema 10.2.1,, es pot demostrar que de fet té mesura exterior 1 el conjunt de trajectòries Hölder contínues en tot interval acotat amb exponent $\beta < \alpha/r$:

$$\forall a \in \mathbb{R}, \quad |\omega(t) - \omega(a)| \leqslant c \cdot |t-a|^\beta \quad , \quad \forall t,\ |t-a| < M, \quad \forall M > 0 \quad . \quad \square$$

En base a aquest criteri de Kolmogorov, tenim el resultat següent per a processos de Markov.

10.2.4 Teorema

Sigui p una funció de transició markoviana tal que $\forall T > 0$, $\exists \alpha > 0$, $r > 0$, $K > 0$ complint

$$\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |y-x|^r \, p(s,x;t,dy) \leqslant K|t-s|^{1+\alpha} \quad , \quad \forall 0 \leqslant s < t \leqslant T \quad .$$

Sigui μ una probabilitat en $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$.
Aleshores, existeix una única probabilitat P en $(C, \mathfrak{B}(C))$ tal que el procés coordenat $\{X_t,\ t \in \mathbb{R}^+\}$ és un procés de Markov continu amb funció de transició p i llei inicial μ.
En particular, $\forall s \geqslant 0$, $\forall x \in \mathbb{R}^d$, existeix una única probabilitat $P_{s,x}$ sobre $(C, \mathfrak{B}(C))$ tal que

$$P_{s,x}\{X_r \equiv x,\ \forall 0 \leqslant r \leqslant s\} = 1$$

i

$$P_{s,x}\big(X_t \in B /_{\mathfrak{F}_r(X)}\big) = p(r, X_r; t, B) \quad , \quad \forall s \leqslant r \leqslant t,\ \forall B \in \mathfrak{B}(\mathbb{R}^d) \quad . \quad \square$$

Com a cas particular, suposem que $p(s, x; t, \cdot)$ té densistat

$$f(y) = \frac{1}{(2\pi(t-s))^{d/2}} \exp\left\{\frac{-|y-x|^2}{2(t-s)}\right\} \quad .$$

Aleshores la hipòtesi del teorema anterior se satisfà amb $\alpha = 1$, $r = 4$, $K = d^2 + 2d$. En efecte:

$$\int_{\mathbb{R}^d} |y-x|^4\, p(s, x; t, dy)$$

és $(d^2 + 2d)(t-s)^2$

Obtenim per tant el següent teorema clàssic:

10.2.5 Teorema de Wiener

Per a cada $s \in \mathbb{R}^+$ i cada $x \in \mathbb{R}^d$, existeix una única probabilitat $P_{s,x}$ en $(C, \mathfrak{B}(C))$ tal que

$$P_{s,x}\{X_r \equiv x,\ \forall 0 \leqslant r \leqslant s\} = 1$$

i

$$P_{s,x}\big(X_t \in B /_{\mathfrak{F}_r(X)}\big) = \int_B \frac{1}{(2\pi(t-r))^{d/2}} \exp\left\{\frac{-|y-x|^2}{2(t-r)}\right\} dy \quad ,$$

$\forall s \leqslant r \leqslant t,\ \forall B \in \mathfrak{B}(\mathbb{R}^d) \quad . \quad \square$

Aquesta mesura s'anomena *mesura de Wiener d-dmensional començant en (s, x).*

10.2.6 Exercici

Existeix una única probabilitat en $(C, \mathfrak{B}(C))$ tal que $\forall n \in \mathbb{N}$,

$$\begin{aligned} &P\{\omega(t_1) \in dx_1, \dots, \omega(t_n) \in dx_n\} \\ &= f(0; t_1, x_1) \cdot f(x_1; t_2 - t_1, x_2) \cdot \ldots \cdot f(x_{n-1}; t_n - t_{n-1}, x_n)\, dx_1 \ldots dx_n \quad , \end{aligned}$$

on f és la funció de (3.2.1).
El procés canònic $X_t(\omega) = \omega(t)$ és, sota aquesta llei, un procés de Wiener estàndar començant en zero.

10.3 Criteris de Dynkin–Kinney per a processos markovians

Sigui p una funció de transició markoviana homogènia.
Per a cada $\varepsilon > 0$, definim

$$\alpha_\varepsilon(t) := \sup_{x\in\mathbb{R}^d} \sup_{s\leqslant t} p(x; s, B(x;\varepsilon)^c) \quad , \tag{10.3.1}$$

on $B(x;\varepsilon)$ és la bola de centre x i radi ε.

10.3.1 Teorema

Suposem que $\forall\varepsilon$, $\exists t > 0$ tal que $\alpha_\varepsilon(t) < 1/2$.
Aleshores: Existeix una família markoviana homogènia $\{X_t,\ t \geqslant 0\}$, $\{P_x\}_{x\in\mathbb{R}^d}$ amb funció de transició p i trajectòries càdlàg.

10.3.2 Teorema

Suposem que $\alpha_\varepsilon(t) \xrightarrow[t\to 0]{} 0$, $\forall\varepsilon > 0$.
Aleshores: Existeix una família markoviana homogènia $\{X_t,\ t \geqslant 0\}$, $\{P_x\}_{x\in\mathbb{R}^d}$ amb funció de transició p, tal que:

1) *Té trajectòries càdlàg.*
2) *És uniformement contínua en probabilitat; és a dir,*

$$\forall x \in \mathbb{R}^d,\ \forall\delta > 0, \quad P_x\{|X_s - X_t| \geqslant \delta\} \xrightarrow[|t-s|\to 0]{} 0 \quad .$$

10.3.3 Teorema

Suposem que $\alpha_\varepsilon(t)/t \xrightarrow[t\to 0]{} 0$, $\forall\varepsilon > 0$.
Aleshores: Existeix una família markoviana homogènia $\{X_t,\ t \geqslant 0\}$, $\{P_x\}_{x\in\mathbb{R}^d}$ amb funció de transició p i trajectòries contínues.

10.3.4 Teorema

Suposem que $\sup_x p(x; t, \mathbb{R}^d - \{x\}) \xrightarrow[t\to 0]{} 0$.
Aleshores: Existeix una família markoviana homogènia $\{X_t,\ t \geqslant 0\}$, $\{P_x\}_{x\in\mathbb{R}^d}$ amb funció de transició p i trajectòries esglaonades, càd, amb un nombre finit de salts en cada interval temporal acotat. □

Tots aquests teoremes valen també si l'espai d'estats és un espai polonès.
No puc donar una referència òptima per les demostracions. La referència menys dolenta és Wentzell [12]. També hi l'article original de Kinney [8], amb més coses.

10.3.5 Exercici

No podem esperar, llevat de casos molt particulars, que les trajectòries d'un procés de Markov siguin de classe C^1.

11. Espais d'aplicacions lineals contínues

Aquest capítol comença molt molt abstracte. Però descendeix a nivells bastant concrets al final.

Suposarem que tots els espais vectorials que apareguin són sobre $\mathbb{R}$, però tot val també per altres cossos.

11.1 Espais vectorials topològics

11.1.1 Definició

Un *espai vectorial topològic* (evt) és un espai vectorial E amb una topologia "compatible amb l'estructura d'espai vectorial", en el sentit que:

1) La aplicació

$$\begin{aligned} E \times E &\longrightarrow E \\ (x, y) &\longmapsto x + y \end{aligned}$$

és contínua (amb la topologia producte a $E \times E$).

2) La aplicació

$$\begin{aligned} \mathbb{R} \times E &\longrightarrow E \\ (\lambda, x) &\longmapsto \lambda x \end{aligned}$$

és contínua (amb la topologia producte a $\mathbb{R} \times E$).

11.1.2 Definició

Un espai vectorial topològic s'anomena *localment convex* (evtlc) si existeix un sistema fonamental d'entorns de l'origen format per conjunts convexos.
(De vegades, s'imposa també que la topologia sigui Hausdorff.) $\square$

Recordem que una manera de donar una topologia sobre un conjunt és indicant els entorns de cada punt. Un *sistema fonamental d'entorns* (o *base local*) de un punt és un conjunt d'entorns tal que tot altre entorn inclou un d'aquest conjunt.

En un evt, és suficient donar els entorns (o simplement un sistema fonamental d'entorns) de l'origen per tenir determinada la topologia. Els entorns dels demés punts s'obtenen per trasllació.

11.1.3 Exemple

Tot espai de Banach és un evtlc. □

Pràcticament qualsevol evt que poguem citar sense pensar massa és localment convex. Excepció il·lustre: La convergència en probabilitat correspon a una topologia en $L^0(\Omega, \mathfrak{F}, P)$ que, tot i ser metritzable, no dota aquest espai de l'estructura de evtlc. Això sí, és un evt *localment acotat:* Tot punt té un entorn acotat. Menys és res!

11.2 Aplicacions contínues entre evt

Sigui T un conjunt. Sigui $\mathfrak{S}$ una família de subconjunts de T, filtrant superiorment per la relació d'inclusió.

Sigui F un evt. Sigui $\mathcal{N}$ un sistema fonamental d'entorns de l'origen.

Si $S \in \mathfrak{S}$, $N \in \mathcal{N}$, notem

$$M(S,N) := \{f\colon T \to F \text{ tals que } f(S) \subset N\} \quad .$$

La família

$$\{M(S,N):\ S \in \mathfrak{S},\ N \in \mathcal{N}\}$$

és una base d'entorns de l'origen en F^T (l'espai vectorial de totes les funcions $T \to F$) per una topologia que s'anomena la $\mathfrak{S}$-*topologia* (nom complet: *Topologia de la convergència uniforme sobre conjunts $S \in \mathfrak{S}$).* No depèn del sistema fonamental $\mathcal{N}$ escollit.

Amb aquesta topologia, un subspai $G \subset F^T$ (per exemple, $G = F^T$) és un evt si i només si $\forall f \in G$, $\forall S \in \mathfrak{S}$, f és acotada sobre S (i.e. $f(S)$ és un conjunt acotat de F).

Suposem a més:

T és un espai topològic. $\cup\mathfrak{S}$ és dens en T.

Aleshores: Si G és un subspai vectorial de F^T format per funcions contínues en T i acotades sobre cada $S \in \mathfrak{S}$, llavors, amb la $\mathfrak{S}$-topologia, G és un espai vectorial topològic localment convex.

Suposem a més:

T és un evt, que notarem E. $\mathfrak{S}$ està format per conjunts acotats.

G és $\mathcal{L}(E,F)$, els subspai de F^E de les aplicacions lineals i contínues $E \to F$.

Aleshores: $\mathcal{L}(E,F)$, amb la $\mathfrak{S}$-topologia, és un evtlc.

(Nota: No cal aquí que $\cup\mathfrak{S}$ sigui dens en T. És suficient que $\cup\mathfrak{S}$ sigui *total* en E: Que l'espai vectorial que genera sigui dens en E.)

(Nota: Les aplicacions lineals contínues són automàticament *acotades,* en el sentit que ho són sobre tot conjunt acotat.)

11.2.1 Exemples

1) T conjunt.
   F evt.
   $\mathfrak{S}$ format per tots els subconjunts finits de T.
   Aleshores, F^T amb la $\mathfrak{S}$-topologia és isomorf a l'espai topològic producte de "T còpies" de F.

Com suggereix el "nom complet" de la $\mathfrak{S}$-topologia, aquesta correspon a la convergència uniforme sobre els conjunts de $\mathfrak{S}$. La convergència uniforme sobre conjunts finits no és més que la convergència puntual. Per tant, aquesta és exactament la topologia de la convergència puntual.

2) T espai topològic Hausdorff.
F evtlc.
$\mathfrak{S}$ format per tots els conjunts compactes.
Aleshores, amb la $\mathfrak{S}$-topologia (topologia de la convergència uniforme sobre compactes) l'espai de totes les funcions contínues $T \to F$ és un evtlc.

3) El cas de l'espai dual:
E evtlc.
$F = \mathbb{R}$.
L'espai vectorial $E' = \mathcal{L}(E, \mathbb{R})$ de totes les *formes lineals contínues* sobre E, s'anomena *l'espai dual* de E.
Si $x \in E$, $x' \in E'$, s'acostuma a escriure $\langle x', x\rangle := x'(x)$.

A l'espai dual tenen interès dues topologies essencialment:

a) $\mathfrak{S}$ format pels conjunts finits de E.
La $\mathfrak{S}$-topologia corresponent s'anomena *topologia feble-$\star$*. Com en el cas de l'exemple 1), la convergència relacionada és la convergència puntual:
$\{x'_n\}_{n\in\mathbb{N}} \subset E'$ *convergeix feblement-$\star$* a $x' \in E$ si i només si

$$\lim_{n\to\infty} \langle x'_n, x\rangle = \langle x', x\rangle \quad , \quad \forall x \in E \quad .$$

b) $\mathfrak{S}$ format pels conjunts acotats de E (són aquells pels quals per a tot entorn de l'origen N, existeix $\lambda \in \mathbb{R}$ tal que $A \subset \lambda N$).
La $\mathfrak{S}$-topologia corresponent s'anomena *topologia forta-$\star$*. La convergència relacionada és:
$\{x'_n\}_{n\in\mathbb{N}} \subset E'$ *convergeix fortament-$\star$* a $x' \in E$ si i només si

$$\lim_{n\to\infty} \sup_{x\in S} |\langle x'_n, x\rangle - \langle x', x\rangle| = 0 \quad , \quad \forall S \text{ acotat.}$$

Si E és un espai de Banach, aleshores E' amb la topologia forta-$\star$ és Banach. Es pot normar amb

$$\|x'\| := \sup_{\|x\|=1} |\langle x', x\rangle| \quad .$$

La bola unitat tancada $\{x' \in E' : \|x'\| \leqslant 1\}$ és compacte.

4) El cas de l'espai dual del dual:
La situació interessant és la següent: Partim d'un evt E. Posem en E' la topologia feble-$\star$. Amb aquesta topologia, considerem $(E')'$. Aquest espai s'identifica amb E com a espai vectorial: L'aplicació

$$\begin{aligned} E &\longrightarrow (E')' \\ x &\longmapsto \big(x' \mapsto \langle x', x\rangle\big) \end{aligned}$$

és lineal i bijectiva (isomorfisme d'espais vectorials).
A través d'aquest identificació, podem posar en $E = (E')'$ la topologia feble-$\star$ a partir de la feble-$\star$ de E'. Aquesta s'anomena *topologia feble* (o *topologia afeblida*) de E. Sempre és menys fina que la topologia inicial de E.

La convergència feble és doncs: $\{x_n\}_{n\in\mathbb{N}} \subset E$ convergeix feblement a $x \in E$ si i només si

$$\begin{aligned} &\lim_{n\to\infty} \langle x_n, x'\rangle - \langle x, x'\rangle = 0 \quad , \quad \forall x' \in E' \quad , \\ \Leftrightarrow &\lim_{n\to\infty} \big\langle \big(y' \mapsto \langle y', x_n\rangle\big) - \big(y' \mapsto \langle y', x\rangle\big), x' \big\rangle = 0 \quad , \quad \forall x' \in E' \quad , \\ \Leftrightarrow &\lim_{n\to\infty} \big\langle y' \mapsto \langle y', x_n - x\rangle, x' \big\rangle = 0 \quad , \quad \forall x' \in E' \quad , \\ \Leftrightarrow &\lim_{n\to\infty} \langle x', x_n - x\rangle = 0 \quad , \quad \forall x' \in E' \quad . \end{aligned}$$

Si E és un espai de Hilbert, usant topologies fortes o febles, $(E')'$ és isomorf a E. Deduïm que la bola unitat tancada de un espai de Hilbert és *feblement compacta* (compacta per la topologia feble).

5) El cas dels operadors lineals continus entre evt:
Siguin E, F evt. Considerem l'espai de les aplicacions (també anomenades operadors) lineals i contínues $\mathcal{L}(E, F)$.
La $\mathfrak{S}$-topologia, amb $\mathfrak{S}$ els conjunts acotats de E (topologia de la convergència uniforme sobre acotats) s'anomena simplement *topologia de la convergència acotada*
Considerem ara la $\mathfrak{S}$-topologia, amb $\mathfrak{S}$ els conjunts finits de E, és a dir, la topologia de la convergència puntual. Hi ha dues situacions interessants:
Avís: La nomenclatura és bastant confusiva.

a) Posem en F la seva topologia natural. En tal cas, la $\mathfrak{S}$-topologia que resulta en $\mathcal{L}(E, F)$ s'anomena de vegades *topologia forta.* (!?)
$\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{L}(E, F)$ convergeix a $T \in \mathcal{L}(E, F)$ en aquesta topologia (i.e. *convergeix fortament*) si i només si

$$\lim_{n\to\infty} T_n(x) = T(x) \quad \text{en } F, \forall x \in E \quad .$$

b) Posem en F la topologia afeblida. En tal cas, la $\mathfrak{S}$-topologia que resulta en $\mathcal{L}(E, F)$ s'anomena de vegades *topologia feble.* (!?)
$\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{L}(E, F)$ convergeix a $T \in \mathcal{L}(E, F)$ en aquesta topologia (i.e. *convergeix feblement*) si i només si $\forall x \in E$, $T_n(x) \to T(x)$ en la topologia afeblida de F; o sigui, si i només si

$$\lim_{n\to\infty} \langle x', T_n(x)\rangle = \langle x', T(x)\rangle \quad , \quad \forall x \in E, \forall x' \in F \quad .$$

En el cas particular en què E i F sòn espais de Banach, la topologia de la convergència acotada és normable per

$$\|T\| := \sup_{\|x\|=1} \|T(x)\| = \sup_{x\in E} \frac{\|T(x)\|}{\|x\|} \quad .$$

Amb aquesta norma, $\mathcal{L}(E, F)$ és un espai de Banach. De fet, és un *àlgebra de Banach:* $\|TS\| \leqslant \|T\| \cdot \|S\|$.

11.3 Exemples més concrets

11.3.1 Exemple 1

Considerem l'espai de Banach de les funcions $\mathbb{R} \to \mathbb{R}$ contínues i acotades, amb la topologia de la convergència uniforme. Totes les probabilitats sobre $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ són identificables a elements del dual d'aquest espai, via

$$\langle P, f \rangle := \int_{\mathbb{R}} f(x)\, P(dx) \quad . \tag{11.3.1}$$

En efecte, $\langle P, \cdot \rangle$ és lineal, i és contínua pel Teorema de la Convergència Dominada.
La convergència feble-$\star$ d'una successió de probabilitats $\{P_n\}_{n\in\mathbb{N}}$ a una probabilitat P equival a l'enunciat

$$\int_{\mathbb{R}} f(x)\, P_n(dx) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f(x)\, P(dx) \quad , \quad \forall f \text{ contínua i acotada.}$$

Això és el que habitualment anomenem *convergència feble de probabilitats* (sense l'estrella $\star$). Simplement, perquè als probabilistes ens agrada crear confusió.
La convergència forta-$\star$ de probabilitats equival a l'enunciat

$$\lim_{n\to\infty} \sup_{f\in S} \left| \int_{\mathbb{R}} f(x)\, P_n(dx) - \int_{\mathbb{R}} f(x)\, P(dx) \right| = 0 \quad , \quad \forall S \text{ acotat.}$$

És una convergència massa forta per a l'aplicació que es pretén de la Teoria de la Probabilitat. Per exemple, suposem que $\{a_n\}_{n\in\mathbb{N}}$ és una successió de nombres reals, tots diferents, amb límit a. Sempre podrem trobar una funció contínua f, amb $\|f\|_\infty \leqslant 1$, tal que $f(a_n) = 1$ i $f(a) = -1$. Per tant,

$$\sup_{\|f\|_\infty \leqslant 1} |f(a_n) - f(a)| = 2 \quad .$$

Això demostra que no hi ha convergència forta-$\star$ de les Delta de Dirac δ_{a_n} a δ_a. Aquesta topologia no té en compte per res la topologia de $\mathbb{R}$, i no ens interessa.

11.3.2 Exemple 2

Encara sobre convergència de probabilitats: Per què no partim de l'espai de funcions mesurables i acotades (sense imposar continuïtat), que és un espai que sembla més natural per integrar respecte qualsevol probabilitat? Les probabilitats són elements del seu dual, via la mateixa fórmula (11.3.1).

La convergència feble-$\star$ de probabilitats en aquesta situació equival a l'enunciat

$$\int_{\mathbb{R}} f(x)\, P_n(dx) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f(x)\, P(dx) \quad , \quad \forall f \text{ mesurable i acotada.}$$

És immediat que això implica

$$P_n(A) \xrightarrow[n\to\infty]{} P(A) \quad , \quad \forall A \text{ mesurable.}$$

Igual que a l'exemple anterior amb la convergència forta-$\star$, que x i y estiguin a prop en $\mathbb{R}$ no implica que δ_x i δ_y estiguin a prop en aquesta topologia (que, per cert, no és ni metritzable). No interessa tampoc.

11.3.3 Exemple 3

L'esperança condicionada és un operador lineal *fortament continu* de L^1 en L^1. Amb aquesta expressió el que es vol dir és que l'esperança condicionada és un operador lineal continu quan posem en L^1 la seva topologia natural. En efecte, si $X_n \xrightarrow[n\to\infty]{} X$ en $L^1(\Omega, \mathfrak{F}, P)$, aleshores

$$\mathrm{E}\Big[\big|\,\mathrm{E}\left[{}^{X_n}/_{\mathfrak{G}}\right] - \mathrm{E}\left[{}^{X}/_{\mathfrak{G}}\right]\big|\Big] \leqslant \mathrm{E}\Big[\mathrm{E}\left[{}^{|X_n - X|}/_{\mathfrak{G}}\right]\Big] = \mathrm{E}\left[|X_n - X|\right] \xrightarrow[n\to\infty]{} 0 \quad .$$

També és *feblement continu*. Això vol dir: Continu si hi posem en L^1 la topologia afeblida (tant a la sortida com a l'arribada). Com que el dual de L^1 és L^∞, el que estem dient és que

$$\Big[\ \forall Y \in L^\infty,\ \mathrm{E}[YX_n] \xrightarrow[n\to\infty]{} \mathrm{E}[YX]\ \Big] \implies$$
$$\Big[\ \forall Y \in L^\infty,\ \mathrm{E}\left[Y\,\mathrm{E}\left[{}^{X_n}/_{\mathfrak{G}}\right]\right] \xrightarrow[n\to\infty]{} \mathrm{E}\left[Y\,\mathrm{E}\left[{}^{X}/_{\mathfrak{G}}\right]\right]\ \Big] \quad .$$

En efecte, observem que

$$\mathrm{E}\Big[\mathrm{E}\left[{}^{Y}/_{\mathfrak{G}}\right]\cdot\mathrm{E}\left[{}^{Z}/_{\mathfrak{G}}\right]\Big] = \mathrm{E}\Big[\mathrm{E}\left[{}^{Y\cdot\mathrm{E}\left[{}^{Z}/_{\mathfrak{G}}\right]}/_{\mathfrak{G}}\right]\Big] = \mathrm{E}\Big[Y\cdot\mathrm{E}\left[{}^{Z}/_{\mathfrak{G}}\right]\Big] \quad ,$$

per variables Y i Z qualssevol tals que les esperances tinguin sentit. Suposem $Y \in L^\infty$ i apliquem això dues vegades:

$$\mathrm{E}\Big[Y\cdot\mathrm{E}\left[{}^{X_n - X}/_{\mathfrak{G}}\right]\Big] = \mathrm{E}\Big[\mathrm{E}\left[{}^{Y}/_{\mathfrak{G}}\right]\cdot\mathrm{E}\left[{}^{X_n - X}/_{\mathfrak{G}}\right]\Big] = \mathrm{E}\Big[\mathrm{E}\left[{}^{Y}/_{\mathfrak{G}}\right]\cdot(X_n - X)\Big]$$

$\mathrm{E}\left[{}^{Y}/_{\mathfrak{G}}\right]$ també és de L^∞ i per tant l'última expressió tendeix a zero per hipòtesi.
Aquest càlcul és prou fàcil, però de fet es pot deduir la continuïtat feble de la continuïtat forta en general: Si E i F són dos evt i $T\colon E \to F$ és contínua, aleshores és contínua també si en E i F posem les topologies afeblides. □

Les meves referències preferides sobre espais vectorials topològics són Khoan [7] i Schaefer [10].

12. Espais de Banach

12.1 Espais de Banach

Un *espai de Banach* E és un espai vectorial normat i que, com a espai mètric, amb la distància $d(x,y) := \|x-y\|$, és complet.

Observació: Quan en un evt tota successió de Cauchy és convergent, és diu que l'espai és *seqüencialment complet*. La completesa és una propietat més forta. Però per espais normats, i de fet, per a tot evt amb un sistema fonamental d'entorns de l'origen numerable, ambdues coincideixen.

L'espai dual de E és $E' := \mathcal{L}(E,\mathbb{R})$, l'espai vectorial de totes les formes lineals i contínues en E. E' és un espai de Banach amb la norma

$$\|T\| := \sup_{\|x\|\leqslant 1} |Tx| = \sup_{x\in E} \frac{|Tx|}{\|x\|} \quad .$$

Si no es diu el contrari, quan es parla de E' se suposa que s'hi ha posat la topologia corresponent a aquesta norma. Es diu que $\{T_n\}_{n\in\mathbb{N}} \subset E'$ convergeix a $T \in E'$ si ho fa en aquesta topologia; és a dir, si $\|T_n - T\| \xrightarrow[n\to\infty]{} 0$.

Un altre tipus de convergència important en E' és la següent: Es diu que $\{T_n\}_{n\in\mathbb{N}} \subset E'$ convergeix feblement-$\star$ a $T \in E'$ si

$$\forall x \in E, \quad T_n x \xrightarrow[n\to\infty]{} Tx \quad \text{en } \mathbb{R}.$$

12.2 Operadors acotats en espais de Banach

Si E, F són espais de Banach, l'espai vectorial de tots els operadors lineals continus de E en F, notat $\mathcal{L}(E,F)$, és de Banach amb la norma

$$\|T\| := \sup_{\|x\|\leqslant 1} \|Tx\|_F = \sup_{\|x\|\in E} \frac{\|Tx\|_F}{\|x\|_E} \quad .$$

En el cas $E = F$, l'espai $\mathcal{L}(E, E)$ (endomorfismes continus de E) és una àlgebra de Banach amb la composició, o sigui

$$\|TS\| \leqslant \|T\| \cdot \|S\| \quad .$$

Les aplicacions lineals i contínues també s'anomenen *acotades*. Ho són en el sentit que la imatge d'un conjunt acotat és acotat. Acotat i continu s'han de prendre aquí com a sinònims.

12.3 Operadors no acotats en espais de Banach

Siguin E, F espais de Banach. Sigui $\mathcal{D}$ un subspai de E.

Un *operador lineal* de E en F amb domini $\mathcal{D}$ és una aplicació lineal $\mathcal{D} \to F$.

Un operador lineal T de E en F amb domini $\mathcal{D}(T)$ és *acotat* si $\|Tx\| \leqslant M\|x\|$, per un cert $M > 0$.

Si T_1 i T_2 són operadors lineals de E en F amb dominis respectius $\mathcal{D}_1$ i $\mathcal{D}_2$, amb $\mathcal{D}_1 \subset \mathcal{D}_2$, i $T_1(x) = T_2(x)$, $\forall x \in \mathcal{D}_1$, aleshors es diu que T_2 és una extensió de T_1 (abreviadament, $T_1 \subset T_2$).

La topologia producte en $E \times F$ és normable amb la norma (per exemple)

$$\|(x, y)\| = \|x\|_E + \|y\|_F \quad , \tag{12.3.1}$$

i en resulta un espai de Banach (si E i F són Hilbert, prenent la norma equivalent $\|(x, y)\| = \big(\|x\|_E^2 + \|y\|_F^2\big)^{1/2}$, l'espai $E \times F$ és Hilbert. El producte escalar associat és $\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, x_2 \rangle_E + \langle y_1, y_2 \rangle_F$).

Sigui T un operador lineal de E en F amb domini $\mathcal{D}(T) \subset E$. El *graf de* T és el subspai de $E \times F$

$$g(T) := \{(x, Tx) : \ x \in \mathcal{D}(T)\} \quad .$$

Un operador lineal T de E en F és *tancat* si el seu graf $g(T)$ és un subspai tancat de $E \times F$. Això és equivalent a dir: Si $\{x_n\}_{n \in \mathbb{N}}$ és una successió d'elements de $\mathcal{D}(T)$,

$$\left.\begin{array}{l} x_n \xrightarrow[n\to\infty]{} x \text{ en } E \\ Tx_n \xrightarrow[n\to\infty]{} y \text{ en } F \end{array}\right\} \Longrightarrow x \in \mathcal{D}(T) \text{ i } Tx = y \quad .$$

Si T és un operador lineal de E en F amb domini $\mathcal{D}(T)$, la *norma del graf* en $\mathcal{D}(T)$ és la definida per

$$\|x\|_g = \|x\|_E + \|Tx\|_F$$

(si E i F són Hilbert, la norma equivalent $\|(x, y)\| = \big(\|x\|_E^2 + \|y\|_F^2\big)^{1/2}$, proporciona en $\mathcal{D}(T)$ el producte escalar $\langle (x_1, x_2) \rangle_g = \langle x_1, x_2 \rangle_E + \langle Tx_1, Tx_2 \rangle_F$, que el configura com espai pre-Hilbert (potser no sigui complet)).

Si T és un operador lineal de E en F amb domini $\mathcal{D}(T)$, diem que és *tancable* si té una extensió tancada.

12.3.1 Proposició

Sigui T un operador lineal de E en F amb domini $\mathcal{D}(T)$.

Aleshores, T és tancable si i només si no hi ha cap parell $(0, y)$ amb $y \neq 0$ en $\overline{g(T)}$ (l'adherència del graf).

Dit d'una altra manera, si i només si $\forall \{x_n\}_{n\in\mathbb{N}}$ *successió d'elements de* $\mathcal{D}(T)$*,*

$$\left.\begin{array}{l} x_n \xrightarrow[n\to\infty]{} 0 \text{ en } E \\ Tx_n \xrightarrow[n\to\infty]{} y \text{ en } F \end{array}\right\} \Longrightarrow y = 0 \quad .$$

I en cas afirmatiu, $\overline{g(T)}$ *és el graf de la mínima extensió tancada de* T*.*

Demostració: Veiem que $\overline{g(T)}$ és el graf d'un cert operador lineal: $\overline{g(T)}$ és un subspai de $E\times F$. Si $(x_1,y_1),(x_2,y_2)\in\overline{g(T)}$, aleshores $(x_1,y_1)-(x_2,y_2)=(0,y_1-y_2)\in\overline{g(T)}$, i per hipòtesi $y_1=y_2$. Per tant, $\overline{g(T)}$ és el graf d'una certa aplicació, diem-ne $\bar{T}$, en un cert domini. És lineal, perquè si $\bar{T}(x_1)=y_1$, $\bar{T}(x_2)=y_2$ i $\bar{T}(\alpha x_1+\beta x_2)=z$, tindrem $(x_1,y_1),(x_2,y_2),(\alpha x_1+\beta x_2,z)\in\overline{g(T)}$, i com $\overline{g(T)}$ és subspai vectorial, forçosament $\alpha x_1+\beta x_2=z$.

Per altra banda, $\bar{T}$ serà tancat perquè el seu graf és tancat. I és evident que no pot haver-hi una extensió tancada de T més petita que $\bar{T}$, puix que $\overline{g(T)}$ és el mínim tancat que conté $g(T)$.

La implicació contrària és òbvia: Cap extensió de T pot tenir en el seu graf un parell $(0,y)$, $y\neq 0$, perquè no seria una aplicació ben definida ($(0,0)$ sempre està en el graf). □

En la situació de la proposició anterior, el domini de la mínima extensió tancada $\bar{T}$ serà el subspai de E que és l'adherència de $\mathcal{D}(T)$ respecte la norma del graf:

$$\bar{T}\colon \overline{\mathcal{D}(T)}^{\|\cdot\|_g} \subset E \longrightarrow F \quad ,$$

i, si $x\in\overline{\mathcal{D}(T)}^{\|\cdot\|_g}$ i $\{x_n\}_{n\in\mathbb{N}}$ és una successió d'elements de $\mathcal{D}(T)$ que convergeix a x en la norma del graf (és a dir, $x_n\to x$, $Tx_n\to y$) tindrem que $\bar{T}(x)=y$. O sigui,

$$\bar{T}(x)=\lim_{n\to\infty} T(x_n) \quad , \quad \text{en } F.$$

Es dedueix també del paràgraf anterior que si considerem en $\mathcal{D}(T)$ no la norma de E, sino la norma del graf, aleshores $T\colon\mathcal{D}(T)\to F$ és un operador lineal acotat (en el mateix sentit que hem definit abans, encara que aquí $\mathcal{D}(T)$ no té perquè ser complet):

$$\|x\|_g=\|x\|_E+\|Tx\|_F \implies \|Tx\|_F=\|x\|_g-\|x\|_E\leqslant\|x\|_g \quad .$$

Per tant, si T és tancat, és natural considerar l'espai de Banach $(\mathcal{D}(T),\|\cdot\|_g)$, i l'operador

$$T\colon(\mathcal{D}(T),\|\cdot\|_g)\longrightarrow(F,\|\cdot\|_F) \quad ,$$

que serà lineal i acotat.

Suposem ara que T és un operador lineal acotat amb domini $\mathcal{D}(T)$. Si T és tancat, forçosament el seu domini ha de ser tancat en E. Recíprocament, si $\mathcal{D}(T)$ és tancat en E, aleshores T és tancat, perquè: $\{x_n\}_{n\in\mathbb{N}}\subset\mathcal{D}(T)$, $x_n\xrightarrow[n\to\infty]{}0$ en E, $Tx_n\xrightarrow[n\to\infty]{}y$ en F, impliquen que $x\in\mathcal{D}(T)$ i $\|Tx_n-Tx\|_F\leqslant M\|x_n-x\|_E\xrightarrow[n\to\infty]{}0$, d'on $Tx=y$.

Suposem que T és un operador lineal acotat amb domini $\mathcal{D}(T)$ que no és un tancat de E. Aleshores T es pot estendre a $\overline{\mathcal{D}(T)}$: Si $x\in\overline{\mathcal{D}(T)}$ i $\{x_n\}_{n\in\mathbb{N}}\subset\mathcal{D}(T)$, amb $x_n\xrightarrow[n\to\infty]{}x$, aleshores $\|Tx_n-Tx_m\|_F\leqslant M\|x_n-x_m\|_E$, el que implica que $\{Tx_n\}_{n\in\mathbb{N}}$ és de Cauchy en F, i tindrà límit $y\in F$, independent de $\{x_n\}_{n\in\mathbb{N}}$. Definim $Tx=y$.

Es dedueix en definitiva que: Si T és un operador lineal acotat de E en F amb domini $\mathcal{D}(T)$ i de norma $\|T\|$, aleshores T té una única extensió lineal acotada $\bar{T}$ amb domini $\overline{\mathcal{D}(T)}$ i $\|\bar{T}\|=\|T\|$. L'extensió serà un operador lineal tancat.

12.4 Funcions Banach-valuades de variable real

Si X és un espai topològic, sabem què vol dir la continuïtat de funcions $u: X \to E$. Si X és un espai de Banach com E, també hi ha un càlcul diferencial (càlcul diferencial de Fréchet) que permet parlar de funcions diferenciables $u: X \to E$, anàleg al càlcul diferencial de funcions $\mathbb{R}^n \to \mathbb{R}^m$. Finalment, si $(X, \mathfrak{F}, \mu)$ és un espai de mesura, es pot parlar de mesurabilitat i integració de funcions $u: X \to E$, i també, si μ és una mesura a valors en E es pot parlar de mesurabilitat i integració de funcions $X \to \mathbb{R}$ respecte aquesta μ.

En tot cas, només necessitarem el cas en què $X = \mathbb{R}$. Per tant, particularitzem directament a aquesta situació:

Diem que $u: \mathbb{R} \to E$ és *contínua* en t si $\lim_{h\to 0} u(t+h) = u(t)$ en E (o sigui, si $\|u(t+h) - u(t)\|_E \xrightarrow[h\to 0]{} 0$).

Diem que $u: \mathbb{R} \to E$ és *derivable* en t i la seva derivada és $x \in E$ si

$$\lim_{h\to 0} \frac{u(t+h) - u(t)}{h} = x$$

en E (o sigui, si $\|u(t+h) - u(t) - hx\|_E \xrightarrow[h\to 0]{} 0$). Notació: $x = \frac{du}{dt}(t)$.

Si $u: \mathbb{R} \to E$ és derivable en t, i T és un operador lineal acotat de E en F, aleshores $Tu: \mathbb{R} \to F$ és derivable en t, i

$$\frac{d(Tu)}{dt}(t) = T\Big(\frac{du}{dt}(t)\Big) \quad .$$

Quant a la integral, es pot definir una integral "à la Riemann" o "à la Lebesgue". D'integrals à la Lebesgue n'hi ha diverses. La més simple, encara que no és la més general possible, és la integral de Bochner, que funciona de la manera següent:

Sigui $u: \mathbb{R} \to E$ una funció numerablement valuada, és a dir, que es pot escriure com

$$u(t) = \sum_{k=1}^{\infty} a_k \cdot \mathbf{1}_{A_k} \quad , \quad a_k \in E, \quad A_k \in \mathfrak{B}(\mathbb{R}) \quad .$$

Aleshores, $t \mapsto \|u(t)\|$ és una funció real mesurable, i es diu que u és *Bochner integrable* sii $t \mapsto \|u(t)\|$ és Lebesgue integrable, i en cas afirmatiu es defineix

$$\int_{\mathbb{R}} u(t)\,dt := \sum_{k=1}^{\infty} a_k \cdot \lambda(A_k) \quad ,$$

on λ representa la mesura de Lebesgue. La definició té sentit perquè

$$\sum_{k=1}^{\infty} \|a_k\| \cdot \lambda(A_k) = \int_{\mathbb{R}} \|u(t)\|\,dt < +\infty \quad .$$

Sigui $u: \mathbb{R} \to E$ una funció qualsevol. Diem que u és *Bochner integrable* sii existeix una successió $\{u_n\}_{n\in\mathbb{N}}$ de funcions numerablement valuades tal que $u_n \xrightarrow[n\to\infty]{\text{q.p.t.}} u$ i $\lim_{n\to\infty} \int_{\mathbb{R}} \|u_n(t) - u(t)\|\,dt = 0$, i en cas afirmatiu es defineix

$$\int_{\mathbb{R}} u(t)\,dt := \lim_{n\to\infty} \int_{\mathbb{R}} u_n(t)\,dt \quad .$$

La definició té sentit perquè:

a) El límit existeix i és únic:

$$\begin{aligned}&\| \int_{\mathbb{R}} u_n(t)\,dt - \int_{\mathbb{R}} u_m(t)\,dt\| = \| \int_{\mathbb{R}} (u_n(t) - u_m(t))\,dt\| \\ &\leqslant \int_{\mathbb{R}} \|u_n(t) - u_m(t)\|\,dt \leqslant \int_{\mathbb{R}} \|u_n(t) - u(t)\|\,dt + \int_{\mathbb{R}} \|u(t) - u_m(t)\|\,dt \quad ,\end{aligned}$$

d'on la successió d'integrals és de Cauchy en E i per tant té límit únic.

b) Es comprova que la definició és independent de la successió escollida.

La classe de funcions Bochner integrables es caracteritza fàcilment: Són límits q.p.t. de funcions numerablement valuades tals que $\int_{\mathbb{R}} \|u(t)\|\,dt < \infty$.

Escriurem una llista de propietats de la integral de Bochner. Observeu la similitud amb la integral de Lebesgue de funcions reals. Per les demostracions, més propietats de la integral de Bochner, i una introducció a la integral de Pettis (més general), vegeu Hille–Phillips [4].

1) Una combinació lineal de funcions Bochner integrables és Bochner integrable i la integral és lineal.

2) $\| \int_{\mathbb{R}} u(t)\,dt\| \leqslant \int_{\mathbb{R}} \|u(t)\|\,dt$.

3) Si $\{u_n\}_{n\in\mathbb{N}}$ és una successió de funcions Bochner integrables, i $\lim_{n,m\to\infty} \int_{\mathbb{R}} \|u_n(t) - u_m(t)\|\,dt = 0$, aleshores:

   a) Existeix u Bochner integrable tal que

   $$\int_{\mathbb{R}} \|u(t) - u_n(t)\|\,dt \xrightarrow[n\to\infty]{} 0 \quad .$$

   b) Si v també compleix

   $$\int_{\mathbb{R}} \|v(t) - u_n(t)\|\,dt \xrightarrow[n\to\infty]{} 0 \quad ,$$

   aleshores $u = v$ q.p.t.

   c) $\lim_{n\to\infty} \int_{\mathbb{R}} u_n(t)\,dt = \int_{\mathbb{R}} u(t)\,dt$.

4) L'espai de funcions $\mathbb{R} \to E$ Bochner integrables, amb la norma (un cop identifiquem totes les funcions que són iguals q.p.t.)

$$\|u\| := \int_{\mathbb{R}} \|u(t)\|\,dt$$

és un espai de Banach.

5) Si $\{u_n\}_{n\in\mathbb{N}}$ són Bochner integrables, $u_n \xrightarrow[n\to\infty]{\text{q.p.t.}} u$, i $\|u_n(t)\| \leqslant F(t)$, amb $F \in L^1(\mathbb{R})$, aleshores u és Bochner integrable i

$$\lim_{n\to\infty} \int_{\mathbb{R}} u_n(t)\,dt = \int_{\mathbb{R}} u(t)\,dt \quad .$$

6) Si $T: E \to F$ és un operador lineal continu entre espais de Banach, i $u: \mathbb{R} \to E$ és Bochner integrable, aleshores $t \mapsto T(u(t))$ és Bochner integrable i

$$T\Big[\int_{\mathbb{R}} u(t)\,dt\Big] = \int_{\mathbb{R}} T(u(t))\,dt \quad .$$

7) Hi ha un Teorema de Fubini.

8) Si u és Bochner integrable,

$$\lim_{h\to 0} \int_{\mathbb{R}} \|u(t+h) - u(t)\|\,dt = 0$$

9) Teorema Fonamental del Càlcul: Si u és Bochner integrable,

$$\lim_{h\to 0} \frac{1}{h} \int_t^{t+h} u(s)\,ds = u(t) \quad .$$

10) Regla de Barrow: Si u és derivable i $\frac{du}{dt}(t)$ és contínua, aleshores és Bochner integrable, i

$$\int_a^b u(s)\,ds = u(b) - u(a) \quad .$$

13. Semigrups d'operadors

Com hem vist a l'Apartat 7.3, a cada funció de transició markoviana homogènia podem associar un semigrup d'operadors sobre un cert espai de Banach. Hem vist també algunes coses relacionades amb aquests semigrups en el cas de cadenes de Markov amb espai d'estats finit (Exemple 7.3.2).

Estudiem-ho ara en detall. Parlarem de semigrups generals en espais de Banach, però els exemples provindran de funcions de transició markovianes homogènies.

13.1 Semigrups d'operadors en espais de Banach

13.1.1 Definició

Una família d'operadors lineals acotats $\{T_t,\ t \geqslant 0\}$ d'un espai de Banach en si mateix és un *semigrup d'operadors* si

$$T_{s+t} = T_s T_t \quad , \quad \forall s,t \geqslant 0 \quad . \qquad \square$$

A risc de caure en la pedanteria, justificaré el nom de "semigrup". Un *semigrup abstracte* és un conjunt S amb una operació binària que té la propietat associativa. Si, a més, té la propietat commutativa s'anomena *abelià*. Si X és un conjunt, un *semigrup d'aplicacions* de X és una família d'aplicacions, tancada respecte la composició (que ja és sempre associativa). Un semigrup d'aplicacions $\mathfrak{T}$ és una *realització* del semigrup abstracte S si $\mathfrak{T}$ és la imatge de S per un homomorfisme de semigrups. El que diem doncs a la definició anterior és que $\{T_t, t \geqslant 0\}$ és una realització del semigrup abstracte $(\mathbb{R}^+, +)$. Més (molt més!) sobre la teoria analítica de semigrups en Hille-Phillips [4].

13.2 Semigrup d'operadors associats a una funció de transició homogènia

Sigui $(E, \mathfrak{E})$ un espai d'estats (pensem sempre en $\mathbb{R}$ o $\mathbb{R}^n$, però tot el que segueix és molt general).

Sigui $p(\cdot;\cdot,\cdot)$ una funció de transició markoviana homogènia.

Sigui

$$b\mathfrak{E} := \{f\colon E \to \mathbb{R} \text{ mesurables i acotades}\} \quad .$$

$b\mathfrak{E}$ és un espai de Banach amb la norma del suprem

$$\|f\|_\infty = \sup_{x\in E} |f(x)| \quad .$$

Definim, per a tot $t \in \mathbb{R}^+$, $\forall f \in b\mathfrak{E}$,

$$[T_t f](x) := \int_E f(y)\, p(x;t,dy) \quad . \tag{13.2.1}$$

13.2.1 Proposició

En la situació anterior, es té:

1) $T_0 = Id$.

2) $\forall t$, *T_t és un operador lineal i continu de $b\mathfrak{E}$ en $b\mathfrak{E}$ de norma 1. (Concretament, $T_t 1 = 1$.)*

3) $f \geqslant 0 \Rightarrow T_t f \geqslant 0$.

4) La família $\{T_t\}_{t\in\mathbb{R}^+}$ satisfà la propietat de semigrup

$$T_{t+s} = T_t T_s \quad , \quad \forall t,s \in \mathbb{R}^+ \quad .$$

Demostració:

1) $p(x;0,\cdot)$ és δ_x, i per tant $[T_0 f](x) = f(x)$.

2) La linealitat és conseqüència de la linealitat de la integral. Per altra banda, si $\|f\|_\infty \leqslant 1$,

$$\|\int_E f(y)\, p(x;t,dy)\|_\infty \leqslant \sup_{x\in E} \int_E |f(y)|\, p(x;t,dy) \leqslant \sup_{x\in E} \int_E 1\, p(x;t,dy) = 1 \quad ,$$

el que demostra que T_t és continu i de norma més petita o igual a 1.
Finalment, prenent $f \equiv 1$,

$$T_t 1 = \int_E 1\, p(x;t,dy) = 1 \quad ,$$

i per tant la norma és exactament 1.

3) Conseqüència de la positivitat de la integral.

4) Això és de fet l'equació de Chapman–Kolmogorov traduïda als operadors T_t:

$$[T_t(T_s f)](x) = \int_E \int_E f(z)\, p(y;s,dz)\, p(x;t,dy) = \int_E f(z)\, p(x;t+s,dz) = [T_{t+s} f](x) \quad .$$

13.2.2 Observació

Si $\{X_t,\ t \in \mathbb{R}^+\}$, $\{P_x\}_{x\in E}$ és una família markoviana homogènia, podem expressar (13.2.1) com

$$[T_t f](x) = \mathrm{E}_{P_x}[f(X_t)] \quad .$$

13.3 Generador infinitesimal d'un semigrup

Farem abstracció del fet que el semigrup d'operadors provingui d'una funció de transició mitjançant (13.2.1) i estudiarem els semigrups d'operadors en general. Els nostres examples estaran, però, relacionats amb processos markovians.

13.3.1 Definicions

Tota família $\{T_t\}_{t\in\mathbb{R}^+}$ d'operadors lineals continus en un espai de Banach B satisfent $T_{t+s} = T_t T_s$ s'anomena *semigrup d'operadors* en B.
Si a més $\|T_t\| \leqslant 1$, $\forall t$, s'anomena *semigrup de contraccions*. □

Observem que si coneixem T_t per $t \in [0, \varepsilon]$, per un cert $\varepsilon > 0$, aleshores el coneixem per tot t, per la propietat de semigrup. És lògic intentar trobar "característiques infinitesimals" del semigrup al voltant del zero, i veure què es pot deduir del semigrup a partir d'aquestes característiques.

13.3.2 Definició

Sigui $\{T_t\}_{t\in\mathbb{R}^+}$ un semigrup d'operadors en un espai de Banach B, amb $T_0 = \mathrm{Id}$.
Definim
$$Af = \lim_{t\to 0} \frac{T_t f - f}{t} \quad ,$$
on el límit és en el sentit de la norma de B, per aquelles $f \in B$ tal que el límit existeixi.
Dit d'una altra manera,
$$\lim_{t\to\infty} \| \frac{T_t f - f}{t} - Af\|_B = 0 \quad .$$
A és un operador lineal, però en general no acotat. El seu domini serà un cert subconjunt $\mathcal{D}_A$ de B.
A rep el nom de *generador infinitesimal* (o *operador infinitesimal*) del semigrup $\{T_t\}_{t\in\mathbb{R}}$. □

Observeu que, si imposem $T_0 = \mathrm{Id}$, l'operador A és la derivada en 0 per la dreta de la funció $t \mapsto T_t$, definida respecte la topologia forta en $\mathcal{L}(B, B)$ (vegeu Exemple 11.2.1, 5.a).

Podriem considerar el mateix tipus de límit en les altres topologies importants de $\mathcal{L}(B, B)$? Sí:

1. Topologia de la convergència acotada (la més forta, que configura $\mathcal{L}(B, B)$ com espai normat): Definiriem $A = \lim_{t\to 0} \frac{T_t - \mathrm{Id}}{t}$ en aquesta norma. Si el límit existeix, $A \in \mathcal{L}(B, B)$.
2. Topologia feble (la més feble): Definiriem Af com l'element de B tal que per a tot f' en el dual B', $\lim_{t\to 0}\langle f', \frac{T_t f - f}{t}\rangle = \langle f', Af\rangle$, si existeix.

El generador infinitesimal definit mitjançant la convergència forta és el que interessa més.

Donat un semigrup d'operadors, trobar el domini del generador infinitesimal A és en general difícil. Trobar la forma de A en un cert subdomini sol ser més fàcil.

13.3.3 Exemple (Cadena de Markov amb espai d'estats finit)

Ens referim als exemples 7.3.1 i 7.3.2. En 7.3.1, hem vist que
$$[T_t f](i) = \sum_k f(k) \cdot p_t(i, k) \quad .$$

En el nostre cas, p_t és una matriu $N \times N$ i f és un vector N-dimensional. És a dir,

$$T_t f = p_t \cdot f \quad .$$

El generador infinitesimal serà precisament p_0^+, matriu que haviem anomenat A a l'Exemple 7.3.2. Té per domini l'espai sencer $\mathbb{R}^N$, i determina T_t. Si la cadena tingués espai d'estats numerable, no podriem afirmar cap d'aquestes dues coses.

13.3.4 Exemple (moviment determinista amb velocitat constant)
Suposem que $\{X_t,\ t \in \mathbb{R}^+\}$ representa un moviment amb velocitat constant no aleatori en $\mathbb{R}$. O sigui, $X_t = x_0 + tv$, per certes constants x_0 i $v > 0$.

Funció de transició:

$$p(x; t, \cdot) = \delta_{x+tv} \quad .$$

Semigrup:

$$[T_t f](x) = \int_{\mathbb{R}} f(y)\, p(x; t, dy) = \int_{\mathbb{R}} f(y)\, \delta_{x+tv}(dy) = f(x + tv) \quad .$$

Generador infinitesimal: Per a cada x,

$$\lim_{t\to 0} \frac{T_t f - f}{t}(x) = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} = v\frac{d^+ f}{dx}(x)$$

Sembla, doncs, que de moment $A = v\dfrac{d^+}{dx}$ és el candidat, i que el domini ha d'estar contingut en el conjunt de funcions acotades que tenen derivada per la dreta en tot punt. Hem de imposar que aquest límit per cada x sigui uniforme en x. Quan passa això? Observem primer

$$\begin{aligned}\lim_{t\to 0} \|T_t f - f\|_\infty &\leqslant \lim_{t\to 0} \big(\|T_t f - f - tAf\|_\infty + \|tAf\|_\infty\big)\\ &= \lim_{t\to 0} \|T_t f - f - tAf\|_\infty\\ &= \lim_{t\to 0} t\|\frac{T_t f - f}{t} - Af\|_\infty = 0\end{aligned}$$

(ambdós factors tendeixen a zero, de fet). El que estem veient en el fons és que "la derivabilitat implica la continuïtat".
Per tant, en el nostre cas,

$$\|f(x+vt) - f(x)\|_\infty \xrightarrow[t\to 0^+]{} 0 \quad ,$$

el que implica que f ha de ser uniformement contínua. Aleshores,

$$\frac{f(x+vt) - f(x)}{t}$$

també és una funció uniformement contínua en x (i acotada), per a tot t. De on $\dfrac{d^+ f}{dx}$ és el límit uniforme de funcions uniformement contínues i acotades i per tant també és uniformement contínua i acotada. Però tota funció contínua amb derivada per la dreta contínua és derivable.

En resum, hem obtingut que si $f \in \mathcal{D}_A$, aleshores f ha de ser uniformement contínua, acotada, i amb derivada uniformement contínua i acotada. Notem $C^1_{\mathrm{b,unif}}(\mathbb{R})$ aquesta classe de funcions. Tenim doncs $\mathcal{D}_A \subset C^1_{\mathrm{b,unif}}(\mathbb{R})$.
Per altra banda, si $f \in C^1_{\mathrm{b,unif}}(\mathbb{R})$,

$$\begin{aligned}&\lim_{t\to 0}\sup_{x\in\mathbb{R}}\Big|\frac{f(x+vt)-f(x)}{t}-vf'(x)\Big| \\ &=\lim_{t\to 0}\sup_{x\in\mathbb{R}}\Big|v\cdot\frac{f'(x+\lambda_x vt)-f(x)}{t}-vf'(x)\Big| \quad , \quad \text{amb } 0\leqslant\lambda_x\leqslant 1 \quad ,\end{aligned}$$

que és igual a zero, per la continuïtat uniforme de f'.
Conclusió: $\mathcal{D}_A = C^1_{\mathrm{b,unif}}(\mathbb{R})$ i $Af = vf'$.

13.3.5 Exemple (Equació diferencial definida per un camp de vectors)
Sigui $b\colon \mathbb{R}^d \to \mathbb{R}^d$ de classe C^1 amb derivades acotades.
Considerem

$$\begin{cases}\dfrac{dX_t}{dt} = b(X_t)\\ X_0 = x\end{cases}$$

Això defineix un procés determinista.

Funció de transició:

$$p(x;t,\cdot) = \delta_{X_t(x)} \quad .$$

Semigrup:

$$[T_t f](x) = \int_{\mathbb{R}} f(y)\, p(x;t,dy) = \int_{\mathbb{R}} f(y)\, \delta_{X_t(x)}(dy) = f(X_t(x)) \quad .$$

Generador infinitesimal: Sigui $C^1_{\mathrm{comp}}(\mathbb{R}^d)$ l'espai de funcions de classe C^1 a suport compacte. Si $f \in C^1_{\mathrm{comp}}(\mathbb{R}^d)$,

$$f(y) - f(x) = \sum_{j=1}^d \partial_j f(x)\cdot(y_j - x_j) + o(|y-x|) \quad ,$$

i per tant

$$\begin{aligned}&\frac{T_t f(x) - f(x)}{t}\\ &= \frac{f(X_t(x)) - f(x)}{t} = \sum_{j=1}^d \partial_j f(x)\cdot\frac{X_t(x)_j - x_j}{t} + \frac{o(|X_t(x)-x|)}{t} \quad .\end{aligned}$$

Fent $t \to 0$,

$$\lim_{t\to 0}\frac{T_t f(x) - f(x)}{t} = \sum_{j=1}^d b_j(x)\partial_j f(x) \quad ,$$

per a tot x. Anem a veure que el límit és uniforme. Per no complicar la notació, fem-ho suposant $d = 1$. En dimensió superior és idèntic.

Pel Teorema del Valor Mitjà,

$$
\begin{aligned}
&\sup_x \Big| \frac{f(X_t(x)) - f(x)}{t} - f'(x)b(x) \Big| = \\
&= \sup_x \Big| f'(\xi) \cdot \frac{X_t(x) - x}{t} - f'(x)b(x) \Big| \quad ,
\end{aligned}
$$

per un cert ξ en el segment que uneix x i $X_t(x)$.
Com f és de suport compacte (i per tant la seva derivada també), l'anterior valor absolut és zero fora d'aquest compacte. Podem suposar que calculem el suprem en un compacte.
Observem també que:

$$
\begin{aligned}
|X_t(x) - x| &\leqslant \int_0^t |b(X_s(x))| \, ds \\
&\leqslant \int_0^t |b(X_s(x)) - b(x)| \, ds + t|b(x)| \\
&\leqslant C \int_0^t |X_s(x) - x| \, ds + t|b(x)| \quad .
\end{aligned}
$$

Per la desigualtat de Gronwall,

$$
|X_t(x) - x| \leqslant t|b(x)| \cdot \exp\{Ct\} \quad ,
$$

i veiem d'aquí que si fem variar x en un compacte i t en un interval $[0,T]$, aleshores $X_t(x)$ també pren valors en un compacte.

Ara,

$$
\begin{aligned}
&\sup_x \Big| f'(\xi) \cdot \frac{X_t(x) - x}{t} - f'(x)b(x) \Big| \\
&= \sup_x \Big| f'(x) \cdot \Big(\frac{X_t(x) - x}{t} - b(x) \Big) + (f'(\xi) - f'(x)) \frac{X_t(x) - x}{t} \Big| \\
&\leqslant \sup |f'| \sup_x \Big| \frac{X_t(x) - x}{t} - b(x) \Big| + \sup_x |f'(\xi) - f'(x)| \sup_x \Big| \frac{X_t(x) - x}{t} \Big|
\end{aligned}
$$

El primer sumand:

$$
\begin{aligned}
&\sup_x \Big| \frac{X_t(x) - x}{t} - b(x) \Big| \\
&= \sup_x \Big| \frac{1}{t} \int_0^t \big(b(X_s(x)) - b(x) \big) \, ds \Big| \\
&\leqslant \frac{1}{t} \int_0^t \sup_x |b(X_s(x)) - b(x)| \, ds \\
&\leqslant \frac{1}{t} K \int_0^t \sup_x |X_s(x) - x| \, ds \\
&\leqslant \frac{1}{t} K \int_0^t \sup_x t|b(x)| \cdot \exp\{Ct\} \, ds \quad ,
\end{aligned}
$$

que té límit zero quan $t \to 0$.
En el segon sumand tenim un factor acotat,

$$
\sup_x \Big| \frac{X_t(x) - x}{t} \Big| = \sup_x \Big| \frac{1}{t} \int_0^t b(X_s(x)) \, ds \Big| \leqslant \sup |b|
$$

(el suprem és sobre un compacte i per tant finit), i un altre factor que té limit zero quan $t \to 0$:

$$\lim_{t\to 0} \sup_x |f'(\xi) - f'(x)| = 0 \quad ,$$

perquè f' és uniformement contínua.

En resum, hem obtingut que $C^1_{\mathrm{comp}}(\mathbb{R}^d) \subset \mathcal{D}_A$ i que

$$[Af](x) = \sum_{j=1}^{d} b_j(x)\partial_j f(x) \quad .$$

Qui és exactament $\mathcal{D}_A$? No ho sé.

13.3.6 Exercici (procés de Wiener)

L'espai $C^2_{b,\mathrm{unif}}$ de funcions de classe C^2 amb derivades fins a ordre 2 contínues i fitades pertany al domini del generador infinitesimal del procés de Wiener (tant en dimensió $d = 1$ com per $d > 1$). Per $f \in C^2_{b,\mathrm{unif}}$, el generador infinitesimal del procés de Wiener en $\mathbb{R}$ és

$$Af(x) = \frac{1}{2} f''(x) \quad ,$$

i el generador infinitesimal del procés de Wiener en $\mathbb{R}^d$ és

$$Af(x) = \frac{1}{2}\Delta f(x) \quad ,$$

on Δ és l'operador de Laplace.
Trobarem els dominis més endavant (Exemple 13.7.3).

13.4 Continuïtat forta del semigrup

A l'Exemple 13.3.4 hem vist que $f \in \mathcal{D}_A$ implica la continuïtat en zero de la funció $\mathbb{R}^+ \to B$ definida per $t \mapsto T_t f$:

$$\Big[\exists Af \in B: \ \|\frac{T_t f - f}{t} - Af\|_B \xrightarrow[t\to 0]{} 0\Big] \quad \Longrightarrow \quad \|T_t f - f\|_B \xrightarrow[t\to 0]{} 0 \quad .$$

Notarem

$$B_0 := \{f \in B: \ \|T_t f - f\|_B \xrightarrow[t\to 0]{} 0\} \quad .$$

Sabem ja que $\mathcal{D}_A \subset B_0 \subset B$.

13.4.1 Proposició

Sigui $\{T_t,\ t \in \mathbb{R}^+\}$ un semigrup d'operadors sobre l'espai de Banach B, amb $T_0 = Id$. Sigui B_0 com abans.
Aleshores:

1) *B_0 és un subspai vectorial tancat de B.*
2) *$T_t(B_0) \subset B_0$, $\forall t$.*
3) *La funció*

$$\begin{aligned} \mathbb{R}^+ &\longrightarrow \mathcal{L}(B_0, B_0) \\ t &\longmapsto T_t \end{aligned}$$

és fortament contínua a l'origen. (La topologia en B_0 és la induïda per B.)

4) *La funció de 3) és de fet fortament uniformement contínua en tot t.*
5) *$\forall f \in \mathcal{D}_A$, $Af \in B_0$.*

Demostració:

1) És fàcil que és subspai. És tancat: Sigui $\{f_n\}_{n\in\mathbb{N}} \subset B_0$ amb $\lim_{n\to\infty} f_n = f$ en B. Tenim que

$$\|T_t f - f\| \leqslant \|T_t f - T_t f_n\| + \|T_t f_n - f_n\| + \|f_n - f\| \leqslant \|T_t\|\|f_n - f\| + \|T_t f_n - f_n\| \quad .$$

Fixem $\varepsilon > 0$ tal que $\|f_n - f\| < \varepsilon$; després agafem $\delta > 0$ tal que $\|T_t f_n - f_n\| < \varepsilon$ i ja està.

3) Obvi.

2) i 4) Suposem $s \leqslant t$.

$$\|T_t f - T_s f\|_B = \|T_s(T_{t-s} f - f)\|_B \leqslant \|T_t\|_{\mathcal{L}(B_0,B_0)} \|T_{t-s} f - f)\|_B \xrightarrow[t-s\to 0]{} 0 \quad ,$$

i d'aquí surten immediatament ambdues afirmacions.

5) Es dedueix de la fórmula de A i el fet que B_0 és tancat.

13.4.2 Observació

De vegades, abans de començar a parlar de generadors infinitesimals, es posa la hipòtesi que el semigrup sigui fortament continu a l'origen ($B_0 = B$). Això no té cap importància especial, perquè sempre ens mourem dins de B_0. En l'aplicació que tenim en ment (semigrups determinats per funcions de transició), vol dir que el semigrup no el definiriem sobre $b\mathfrak{E}$, sino directament en el seu espai de continuïtat forta $B_0 \subset b\mathfrak{E}$.
Per altra banda, si es demana continuïtat forta a l'origen, no cal posar $T_0 = \mathrm{Id}$, perquè es dedueix: Per $f \in B_0$,

$$T_0 f = T_0(\lim_{t\to 0} T_t f) = \lim_{t\to 0} T_0 T_t f = \lim_{t\to 0} T_{0+t} f = f \quad .$$

Tot plegat no és res profund. És qüestió de gustos.

13.4.3 Proposició

Sigui $f \in \mathcal{D}_A$.
Aleshores, $T_t f \in \mathcal{D}_A$ per a tot t, la funció $t \mapsto T_t f$ és derivable, i

$$\frac{d}{dt} T_t f = A T_t f = T_t A f \quad , \tag{13.4.1}$$

que també es pot expressar

$$T_t f - f = \int_0^t A T_s f \, ds = \int_0^t T_s A f \, ds \quad . \tag{13.4.2}$$

Demostració:
1) $T_t f \in \mathcal{D}_A$ i $A T_t f = T_t A f$:

$$\begin{aligned} \Big\|\frac{T_h T_t f - T_t f}{h} - T_t A f\Big\|_B &= \Big\|T_t\Big(\frac{T_h f - f}{h} - A f\Big)\Big\|_B \\ &\leqslant \|T_t\| \cdot \Big\|\frac{T_h f - f}{h} - A f\Big\|_B \xrightarrow[h\to 0]{} 0 \quad . \end{aligned}$$

2) $A T_t f = \frac{d}{dt} T_t f$:

Sabem que $T_t f \in \mathcal{D}_A$. Llavors,

$$AT_t f = \lim_{h \to 0} \frac{T_h T_t f - T_t f}{h} = \lim_{h \to 0} \frac{T_{t+h} f - T_t f}{h} = \frac{d^+}{dt} T_t f \quad .$$

Només queda veure que, per $t > 0$, la derivada per l'esqurra $\frac{d^-}{dt} T_t f$ també existeix i coincideix amb $T_t A f$:

$$\begin{aligned} &\| \frac{T_{t-h} f - T_t f}{-h} - T_t A f \|_B = \| T_{t-h} \Big(\frac{f - T_h f}{-h} \Big) - T_h A f \|_B \\ &\leqslant \| T_{t-h} \Big(\frac{f - T_h f}{-h} \Big) - T_{t-h} A f \|_B + \| T_{t-h} A f - T_t A f \|_B \quad . \end{aligned}$$

El primer sumand és menor o igual que

$$\| T_{t-h} \| \cdot \| \frac{T_h f - f}{h} - A f \|_B \xrightarrow[h \searrow 0]{} 0 \quad .$$

El segon també tendeix a zero, perquè $Af \in B_0$ (Proposició 13.4.1).

3) La fórmula (13.4.2) és la forma integral de (13.4.1).

13.4.4 Observació

El resultat anterior es pot formular dient que $u_t = T_t f$ és solució de l'equació diferencial en espais de Banach

$$\begin{cases} \dfrac{du_t}{dt} = Au \\ u_0 = f \end{cases}$$

13.5 La resolvent del semigrup

13.5.1 Proposició

Sigui T_t un semigrup de contraccions en un espai de Banach B.
Sigui A el seu generador infinitesimal.
Sigui $g \in B_0$.
Aleshores, per a tot $\lambda > 0$, l'equació

$$\lambda f - A f = g \tag{13.5.1}$$

té una única solució $f \in \mathcal{D}_A$, i ve donada per

$$f = R_\lambda g := \int_0^\infty e^{-\lambda t} T_t g \, dt$$

A més, l'operador $R_\lambda \colon B_0 \to B_0$ és lineal, continu, amb $\|R_\lambda\| \leqslant 1/\lambda$.

Demostració: La integral té sentit, perquè la norma de l'integrand està acotada per la funció integrable $e^{-\lambda t}\|g\|$.

De la propietat 2) de la integral de Bochner (Apartat 12.4), es té

$$\|R_\lambda g\| \leqslant \int_0^\infty e^{-\lambda t} \|T_t g\| \, dt \leqslant \int_0^\infty e^{-\lambda t} \|g\| \, dt = \frac{1}{\lambda} \|g\| \quad .$$

Anem a veure que $f = R_\lambda g$ és de $\mathcal{D}_A$ i solució de l'equació de l'enunciat:

$$\begin{aligned} T_h f = T_h R_\lambda g &= \int_0^\infty e^{-\lambda t} T_{t+h} g\, dt = \int_h^\infty e^{-\lambda(t-h)} T_t g\, dt \\ &= e^{\lambda h} \int_h^\infty e^{-\lambda t} T_t g\, dt = e^{\lambda h} \Big[\int_0^\infty e^{-\lambda t} T_t g\, dt - \int_0^h e^{-\lambda t} T_t g\, dt \Big] \\ &= e^{\lambda h} \Big[f - \int_0^h e^{-\lambda t} T_t g\, dt \Big] \quad . \end{aligned}$$

Per tant,

$$\frac{T_h f - f}{h} = \frac{e^{\lambda h} - 1}{h} - \frac{e^{\lambda h}}{h} \int_0^h e^{-\lambda t} T_t f\, dt \quad .$$

Fent $h \to 0$, veiem que el límit existeix i que és $\lambda f - g$, de on $f \in \mathcal{D}_A$ i satisfà l'equació (13.5.1).

Anem a veure la unicitat de la solució. Si hi ha dues solucions diferents, la seva resta φ serà un element no nul de $\mathcal{D}_A$ tal que $\lambda\varphi - A\varphi = 0$. Per (13.4.1), tindrem

$$\frac{dT_t\varphi}{dt} = \lambda T_t \varphi \quad .$$

Intuïtivament, la solució d'aquesta equació diferencial, tenint en compte la condició inicial $T_0\varphi = \varphi$, és

$$T_t \varphi = e^{\lambda t} \varphi \quad ,$$

d'on $\|\varphi\| = e^{-\lambda t} \|T_t \varphi\| \leqslant e^{-\lambda t} \|\varphi\|$, cosa que només compleix $\varphi = 0$. Fem-ho més rigorós:

$$\begin{aligned} \frac{d(e^{-\lambda t} T_t \varphi)}{dt} &= e^{-\lambda t} \frac{d(T_t\varphi)}{dt} - \lambda e^{-\lambda t} T_t \varphi \\ &= e^{-\lambda t} \lambda T_t \varphi - \lambda e^{-\lambda t} T_t \varphi = 0 \quad . \end{aligned}$$

Integrant entre 0 i ∞ ens queda immediatament $\varphi = 0$. □

El que diu aquesta proposició, en altres paraules, és que, per qualsevol $\lambda > 0$, l'operador $\lambda \mathrm{Id} - A$ és bijectiu de $\mathcal{D}_A$ en B_0, i es defineix $R_\lambda := (\lambda \mathrm{Id} - A)^{-1}$.

13.5.2 Definició

L'operador bijectiu $R_\lambda \colon B_0 \to \mathcal{D}_A$ s'anomena la *resolvent* de l'operador A (i també del semigrup $\{T_t\}_{t \in \mathbb{R}^+}$). □

També es pot definir directament $R_\lambda g := \displaystyle\int_0^\infty e^{-\lambda t} T_t g\, dt$, per $g \in B$ tal que la integral estigui ben definida. Això no té importància.

13.6 El potencial

Si posem $\lambda = 0$ a la definició de la resolvent, obtenim un operador que en general és no acotat, encara que el semigrup sigui de contraccions:

$$Rg := \int_0^\infty T_t g\, dt$$

s'anomena *potencial* del semigrup T_t. El domini de R esta format per les funcions g tals que $T_t g$ és integrable.

Posant $\lambda = 0$ en la Proposició 13.5.1, obtenim el següent resultat pel potencial:

13.6.1 Proposició

Sigui $\{T_t,\ t \in \mathbb{R}^+\}$ *un semigrup de contraccions en un espai de Banach* B.
Sigui A *el seu generador infinitesimal i* R *el potencial.*
Sigui $g \in B_0$.
Aleshores:

1) La funció $f := Rg$ *pertany a* $\mathcal{D}_A$ *i és solució de l'equació*

$$-Af = g \quad . \tag{13.6.1}$$

2) Si a més l'operador R *és acotat,* $f := Rg$ *és la única solució de (13.6.1), el que ens diu que en aquest cas* A *és bijectiu de* $\mathcal{D}_A$ *en* B_0*, i* $R = A^{-1}$*.* □

Observeu que la resolvent R_λ d'un semigrup $\{T_t,\ t \in \mathbb{R}^+\}$ coincideix amb el potencial del semigrup $\{e^{-\lambda t}T_t,\ t \in \mathbb{R}^+\}$. D'aquí que es poden estudiar les propietats de les resolvents a partir de les propietats dels potencials.

13.7 Relacions entre semigrup, generador infinitesimal i resolvent

Ens fem dues preguntes:

a) El generador infinitesimal determina el semigrup? El que tenim fins ara és que

$$T_t \longrightarrow A \longleftrightarrow R_\lambda \quad ,$$

on $\longrightarrow$ vol dir 'determina'.

b) Quins operadors A en B_0 són generadors infinitesimals d'algun semigrup d'operadors?

Fem primer la pregunta a). La resposta està en el Teorema 13.7.2. Veurem abans:

13.7.1 Proposició

Sigui $\{T_t,\ t \in \mathbb{R}^+\}$ *un semigrup de contraccions fortament continu en un espai de Banach* B_0.
Sigui A *el seu generador infinitesimal.*
Aleshores:

1) $\overline{\mathcal{D}_A} = B_0$ *(es diu que* A *està densament definit en* B_0*).*

2) L'operador A *és tancat.*

Demostració:
1) Si $f \in B_0$, construïrem una successió d'elements de $\mathcal{D}_A$ que convergeixen a f. Recordem que $\mathcal{D}_A = R_\lambda(B_0)$, $\forall\lambda$. La successió que anirà bé és $\lambda R_\lambda f$, fent variar λ cap a infinit:

$$\begin{aligned}
\|\lambda R_\lambda f - f\| &= \|\int_0^\infty \lambda e^{-\lambda t}(T_t f - f)\,dt\| \\
&\leqslant \|\int_0^h \lambda e^{-\lambda t}(T_t f - f)\,dt\| + \|\int_h^\infty \lambda e^{-\lambda t}(T_t f - f)\,dt\| \\
&\leqslant \sup_{t\leqslant h}\|T_t f - f\| + \int_h^\infty \lambda e^{-\lambda t}(\|T_t f\| + \|f\|)\,dt \\
&\leqslant \sup_{t\leqslant h}\|T_t f - f\| + 2\|f\|\int_0^h \lambda e^{-\lambda t}\,dt \xrightarrow[\lambda\to\infty]{} \sup_{t\leqslant h}\|T_t f - f\| \quad ,
\end{aligned}$$

per Convergència Dominada. Com que el semigrup és fortament continu en B_0, fent ara $h \to 0$, obtenim la convergència buscada.

2) Sigui $\{f_n\}_{n\in\mathbb{N}}$ una successió d'elements de $\mathcal{D}_A$ amb límit f, tal que Af_n té límit g. Hem de veure que $f \in \mathcal{D}_A$ i $Af = g$.

$$f_n = R_\lambda \circ (\lambda \mathrm{Id} - A)(f_n) = R_\lambda(\lambda f_n - Af_n) \quad .$$

Fent $n \to \infty$ en aquesta igualtat, tenint en compte que R_λ és continu,

$$f = R_\lambda(\lambda f - g) \quad , \tag{13.7.1}$$

que implica que f és de la imatge de R_λ, que és $\mathcal{D}_A$.
Per altra banda, aplicant $R_\lambda^{-1} = \lambda \mathrm{Id} - A$ en (13.7.1), obtenim

$$(\lambda \mathrm{Id} - A)f = \lambda f - g \quad ,$$

d'on $Af = g$.

13.7.2 Teorema

Si dos semigrups de contraccions $\{T_t,\ t \in \mathbb{R}^+\}$ i $\{\tilde{T}_t,\ t \in \mathbb{R}^+\}$ tenen la mateixa resolvent (equivalentment, el mateix generador infinitesimal) aleshores $T_t = \tilde{T}_t$, $\forall t$ sobre B_0.

Demostració: Recordem que $\overline{\mathcal{D}_A} = B_0$, i per tant l'espai B_0 està determinat pel generador infinitesimal (equivalentment, per la resolvent), i ha de ser el mateix pels dos semigrups.

Que les resolvents coincideixin vol dir que $\forall f \in B_0$, $\forall \lambda > 0$,

$$\int_0^\infty e^{-\lambda t} T_t f \, dt = \int_0^\infty e^{-\lambda t} \tilde{T}_t f \, dt \quad .$$

Però això implica $T_t f = \tilde{T}_t f$, per a tot t. Vegeu Dynkin [3], I.3.1.6, pels detalls. Essencialment, les integrals anteriors són transformades de Laplace, per la qual hi ha una fórmula d'inversió, anàlogament al cas de funcions valuades en $\mathbb{R}$. □

Atenció: T_t queda determinat sobre B_0. No sobre B. Per tant, no podem treure d'aquí la conclusió que si T_t és el semigrup associat a una funció de transició markoviana, aquesta quedi determinada.

13.7.3 Exemple

Amb l'ajuda de la resolvent, calcularem el domini del generador inifinitesimal associat al procés de Wiener en dimensió $d = 1$.
1r pas: $B_0 = C_{b,\mathrm{unif}}$:
Veiem primer que per a tot $f \in B$ (l'espai de funcions reals mesurables acotades), $R_\lambda f \in C_{b,\mathrm{unif}}$ (l'espai de funcions uniformement contínues i acotades):
Sabem que

$$T_t f(x) = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} f(y) e^{-(x-y)^2/2t} \, dy \quad .$$

La resolvent en f serà

$$\begin{aligned} R_\lambda f(x) &= \int_0^\infty e^{-\lambda t} \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} f(y) e^{-(x-y)^2/2t} \, dy \, dt \\ &= \int_{\mathbb{R}} f(y) \int_0^\infty e^{-\lambda t} \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t} \, dt \, dy \\ &= \frac{1}{\sqrt{2\lambda}} \int_{\mathbb{R}} f(y) e^{-\sqrt{2\lambda}|x-y|} \, dy \end{aligned}$$

(la última igualtat no és fàcil), d'on es veu sense dificultat que $R_\lambda f \in C_{b,\text{unif}}$.
Com que $\mathcal{D}_A = R_\lambda(B_0) \subset C_{b,\text{unif}}$, i $C_{b,\text{unif}}$ és un subspai tancat de B, tenim que $B_0 = \overline{\mathcal{D}_A} \subset C_{b,\text{unif}}$.
Per altra banda, també és un exercici fàcil comprovar que $B_0 \supset C_{b,\text{unif}}$.
Conclusió: $B_0 = C_{b,\text{unif}}$.

2n pas: Veurem ara que $\mathcal{D}_A \subset C^2_{b,\text{unif}}$:
Com que $R_\lambda(B_0) = \mathcal{D}_A$ i $B_0 = C_{b,\text{unif}}$, només hem de comprovar que

$$\forall f \in C_{b,\text{unif}}, \quad R_\lambda f \in C^2_{b,\text{unif}} \quad .$$

I això es veu directament:

$$\begin{aligned} R_\lambda f(x) &= \frac{1}{\sqrt{2\lambda}} \int_{\mathbb{R}} f(y) e^{-\sqrt{2\lambda}|x-y|}\, dy \\ &= \frac{1}{\sqrt{2\lambda}} \int_{-\infty}^{x} f(y) e^{-\sqrt{2\lambda}(x-y)}\, dy + \frac{1}{\sqrt{2\lambda}} \int_{x}^{\infty} f(y) e^{-\sqrt{2\lambda}(y-x)}\, dy \end{aligned}$$

és derivable respecte de x i

$$\begin{aligned} \frac{d}{dx} R_\lambda f(x) &= \frac{1}{\sqrt{2\lambda}} \Big[f(x) - \sqrt{2\lambda} \int_{-\infty}^{x} f(y) e^{-\sqrt{2\lambda}(x-y)}\, dy - f(x) + \sqrt{2\lambda} \int_{x}^{\infty} f(y) e^{-\sqrt{2\lambda}(y-x)}\, dy \Big] \\ &= -\int_{-\infty}^{x} f(y) e^{-\sqrt{2\lambda}(x-y)}\, dy + \int_{x}^{\infty} f(y) e^{-\sqrt{2\lambda}(y-x)}\, dy \quad , \end{aligned}$$

que torna a ser derivable respecte x i

$$\begin{aligned} \frac{d^2}{dx} R_\lambda f(x) &= -f(x) - \sqrt{2\lambda} \int_{-\infty}^{x} f(y) e^{-\sqrt{2\lambda}(x-y)}\, dy - f(x) + \sqrt{2\lambda} \int_{x}^{\infty} f(y) e^{-\sqrt{2\lambda}(y-x)}\, dy \\ &= -2f(x) - \sqrt{2\lambda} \Big[\int_{-\infty}^{x} f(y) e^{-\sqrt{2\lambda}(x-y)}\, dy - \int_{x}^{\infty} f(y) e^{-\sqrt{2\lambda}(y-x)}\, dy \Big] \quad , \end{aligned}$$

que ja no és derivable. En tot cas,

$$R_\lambda f(x), \quad \frac{d}{dx} R_\lambda f(x), \quad \frac{d^2}{dx} R_\lambda f(x)$$

són uniformement contínues i acotades.

3r pas: Finalment, sabem per càlcul directe (Exercici 13.3.6) que $C^2_{b,\text{unif}} \subset \mathcal{D}_A$. Arribem a la conclusió que $\mathcal{D}_A = C^2_{b,\text{unif}}$ i $Af = \frac{1}{2} f''(x)$.

En el cas del procés de Wiener en $\mathbb{R}^d$, només és té $C^2_{b,\text{unif}} \subset \mathcal{D}_A$ i no hi ha igualtat. Això és conseqüència de què A ha de ser un operador tancat. En canvi, se sap que l'operador de Laplace no és tancat (si $d > 1$). Per tant, el generador infinitesimal té un domini més ampli.

De fet, se sap que $\frac{1}{2}\Delta$ és tancable, i que el generador infinitesimal és la seva mínima extensió tancada. □

Fem ara la pregunta b) del principi de l'apartat. Quins operadors lineals A en un espai de Banach són generadors infinitesimals d'algun semigrup?

Per una banda, si A és acotat, A és generador infinitesimal d'un semigrup, que a més es construeix molt fàcilment a partir de A. Es pot representar com $T_t = e^{tA}$ (a precisar el significat d'això). Per altra banda, si A és no acotat, en general no, però hi ha un conjunt de condicions necessàries i suficients per tal que A doni lloc a un semigrup de contraccions.

Per estudiar el primer cas, definirem primer la funció exponencial d'un operador acotat, i veurem algunes de les seves propietats.

Siguin B un espai de Banach. Considerem en $\mathcal{L}(B,B)$ la topologia de la convergència acotada, que és normable i el configura com a espai de Banach.

Sobre qualsevol espai normat es poden considerar sèries, i funcionen molt semblant al cas de les sèries de nombres reals o complexos. Es pot definir la convergència d'una sèrie com la convergència de la successió de sumes parcials.

Si B és un espai normat, es diu que una sèrie $\sum_{n\geqslant 1} f_n$ és *absolutament convergent* sii la sèrie numèrica $\sum_{n\geqslant 1} \|f_n\|$ és convergent. (També es diu *normalment convergent*, si es vol insistir en què tenim una "norma" i no un "valor absolut". Qüestió de gustos.)

13.7.4 Proposició

Sigui B un espai de Banach.
Sigui $\sum_{n\geqslant 1} f_n$ una sèrie en B.
Si $\sum_{n\geqslant 1} f_n$ és absolutament convergent, aleshores és convergent i

$$\|\sum_{n=1}^{\infty} f_n\| \leqslant \sum_{n=1}^{\infty} \|f_n\| \quad .$$

Demostració: La convergència absoluta de la sèrie implica que

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N}:\ m > k \geqslant n_0 \ \Rightarrow\ \sum_{n=k}^{m} \|f_n\| \leqslant \varepsilon \quad .$$

Donat que

$$\|\sum_{n=k}^{m} f_n\| \leqslant \sum_{n=k}^{m} \|f_n\| \quad ,$$

estem dient que la successió de sumes parcials és de Cauchy, i com l'espai és complet, és convergent.

La segona afirmació és la desigualtat triangular passada al límit, usant que la norma és una funció contínua. □

La gràcia de les sèries absolutament convergents és, al igual que per sèries numèriques, que són també *commutativament convergents* (una altra nomenclatura sense importància que hom troba de tant en tant), és a dir, que qualsevol reordenació dels seus termes dóna lloc a una altra sèrie absolutament convergent i amb el mateix límit.

Sigui $A \in \mathcal{L}(B,B)$. Es defineix l'operador l'*exponencial de A* com

$$e^A := \sum_{n=0}^{\infty} \frac{1}{n!} A^n \quad . \tag{13.7.2}$$

Recordem que $A, B \in \mathcal{L}(B,B) \Rightarrow \|AB\| \leqslant \|A\|\|B\|$. Per tant,

$$\|\frac{A^n}{n!}\| \leqslant \frac{\|A\|^n}{n!} \quad , \tag{13.7.3}$$

la sèrie de (13.7.2) és absolutament convergent i la seva suma existeix. $e^A \in \mathcal{L}(B,B)$. Obseveu que si B és de dimensió finita, estem definint la coneguda "exponencial d'una matriu".

13.7.5 Proposició

Propietats de e^A:

1) $\|e^A\| \leqslant e^{\|A\|}$.
2) $e^{c\mathrm{Id}} = e^c Id$.
3) $AB = BA \Rightarrow e^A e^B = e^{A+B}$.
4) $\lim\limits_{t\to 0} \frac{e^{tA}-\mathrm{Id}}{t} = A$ *(en $\mathcal{L}(B,B)$!, o sigui, en la topologia de la convergència acotada).*

Demostració: 1) és (13.7.3) passada al límit. 2) és evident. 3) es comprova multiplicant les sèries (el fet que dues sèries absolutament convergents és puguin multiplicar i en resulti una sèrie absolutament convergent és cert com per a sèries numèriques).
4)

$$\|e^{tA} - \mathrm{Id} - tA\| = \|\sum_{n=2}^{\infty} \frac{A^n}{n!} t^n\| \leqslant \sum_{n=2}^{\infty} \frac{\|A^n\|}{n!} t^n = e^{t\|A\|} - 1 - t\|A\| \xrightarrow[t\to\infty]{} 0 \quad . \qquad \square$$

13.7.6 Corol·lari

Per la Proposició 13.7.5, 3), $\{e^{tA}\}$ és un semigrup d'operadors. Per 4), el seu generador infinitesimal és A.
Per tant: $A \in \mathcal{L}(B,B) \Rightarrow A$ és generador infinitesimal d'un semigrup d'operadors: $T_t = e^{tA}$, $t \in \mathbb{R}^+$. $\square$

De tota manera, això no ens assegura que $T_t = e^{tA}$ sigui un semigrup de contraccions.

Veurem ara el criteri general d'obtenció d'un semigrup de contraccions a partir d'un operador donat A, acotat o no:

13.7.7 Teorema de Hille–Yosida

Sigui B_0 un espai de Banach.
Sigui A un operador lineal de B_0 en B_0 amb domini $\mathcal{D}_A$.
Aleshores, A és el generador infinitesimal d'un semigrup de contraccions fortament continu en B_0 si i només si

a) $\overline{\mathcal{D}_A} = B_0$.
b) $\forall \lambda > 0$, *existeix* $R_\lambda := (\lambda Id - A)^{-1} \colon B_0 \to B_0$ *lineal continu.*
c) $\|R_\lambda\| \leqslant 1/\lambda$. $\square$

Completem el Teorema de Hille-Yosida amb l'enunciat següent:

13.7.8 Teorema

Suposem que B_0 és subspai de l'espai de funcions contínues i acotades B.
Sigui A un operador lineal de B_0 en B_0 complint les condicions a), b), c) del Teorema de Hille-Yosida.

Sigui $\{T_t,\ t \in \mathbb{R}^+\}$ *el semigrup de contraccions fortament continu en* B_0 *donat pel Teorema de Hille-Yosida.*
Aleshores:

1) $\{T_t,\ t \in \mathbb{R}^+\}$ *és positiu* ($f \geqslant 0 \Rightarrow T_t f \geqslant 0$) *si i només si* $(\lambda Id - A)^{-1}$ *és positiu.*

2) $T_t 1 = 1$, *per a tot* t *si i només si* $a \in \mathcal{D}_A$ *i* $A1 = 0$. □

Això ens dóna les condicions addicionals per tal que el semigrup $\{T_t,\ t \in \mathbb{R}^+\}$ sigui l'associat a una funció de transició markoviana homogènia.

Referències

[1] Aureli Alabert. *Mesura i probabilitat*. Publicacions U.A.B., 1996.

[2] Robert Ash. *Real Analysis and Probability*. Academic Press, 1972.

[3] Eugene B. Dynkin. *Markov Processes*. Number 121 in Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen. Springer-Verlag, 1965.

[4] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups*. American Mathematical Society, 1957.

[5] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Number 103 in GTM. Springer–Verlag, 1988.

[6] D. G. Kendall. Introduction to stochastic analysis. In D. G. Kendall and E. F. Harding, editors, *Stochastic Analysis. A Tribute to the Memory of Rollo Davidson*. Wiley, 1973.

[7] Vo-Khac Khoan. *Distributions. Analyse de Fourier. Opérateurs aux dérivées partielles*. Librairie Vuibert, 1972.

[8] J.R. Kinney. Continuity properties of sample functions of markov processes. *Transactions American Mathematical Society*, 74, 1953.

[9] Paul-André Meyer. *Probabilités et potentiel*. Hermann, 1966.

[10] Helmut H. Schaefer. *Topological vector spaces*. Springer-Verlag, 1986.

[11] Daniel Stroock and S.R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Springer-Verlag, 1979.

[12] A.D. Wentzell. *A course in the theory of stochastic processes*. McGraw-Hill, 1981.

[13] David Williams. *Diffusions, Markov Processes, and Martingales (vol. 1)*. Wiley, 1979.

[14] David Williams. *Probability with martingales*. Cambridge University Press, 1991.