Linear inverse problems and their regularisation

Umut Çetin

London School of Economics

Bachelier Workshop, September 2016

An inverse problem for Markov processes

- Suppose X is an **E**-valued Markov process with semi-group $(P_t)_{t\geq 0}$ and infinitesimal generator A, where **E** is some topological space. Given a function g and $T > 0$ does there exists an f so that

$$g(x) = P_T f(x) := E^x[f(X_T)]? \tag{1}$$

An inverse problem for Markov processes

- Suppose X is an **E**-valued Markov process with semi-group $(P_t)_{t \geq 0}$ and infinitesimal generator A, where **E** is some topological space. Given a function g and $T > 0$ does there exists an f so that

$$g(x) = P_T f(x) := E^x[f(X_T)]? \tag{1}$$

- If the answer to the above is affirmative, then we can define a function $u : [0, T] \times E \mapsto \mathbb{R}$ by $u(t, \cdot) = P_{T-t} f$ which will solve

$$\begin{aligned} u_t + Au &= 0; \\ u(0, \cdot) &= g; \end{aligned} \tag{2}$$

a backward PDE with an initial condition!

Backward equations with an initial condition

- A class of ill-posed PDEs akin to the one given in (2) has been recently studied in Mathematical Finance literature in the context of forward utilities (see, e.g., the works by Nadtochiy as well as Tehranchi and Shkolnikov et al.) .
- However, these works search for solutions of (2) over $[0, \infty)$ while the linear inverse problem (1) is related to the L^2-solutions over a bounded interval $[0, T]$.

Backward equations with an initial condition

- A class of ill-posed PDEs akin to the one given in (2) has been recently studied in Mathematical Finance literature in the context of forward utilities (see, e.g., the works by Nadtochiy as well as Tehranchi and Shkolnikov et al.) .
- However, these works search for solutions of (2) over $[0, \infty)$ while the linear inverse problem (1) is related to the L^2-solutions over a bounded interval $[0, T]$.
- This is a crucial distinction as a solution over $[0, \infty)$ necessarily require the initial condition, g, to be unbounded while the inverse problem (1) with a solution in L^2 implicitly imposes boundary conditions on g given by the behaviour of the diffusion. In particular, g must vanish at the natural or absorbing boundaries.

- If g is a density then the linear inverse problem with positivity constraint is the answer to the following question:
  Given the European option price data at time-T can we find an initial *distribution for our* favourite *diffusion process to obtain a model consistent with the observed option prices?*
- As such, the solutions of (1) provide alternative “answers" to certain questions in Mathematical Finance that are traditionally attacked using Skorokhod embedding.

- If g is a density then the linear inverse problem with positivity constraint is the answer to the following question:

  Given the European option price data at time-T can we find an initial *distribution for our* favourite *diffusion process to obtain a model consistent with the observed option prices?*
- As such, the solutions of (1) provide alternative “answers" to certain questions in Mathematical Finance that are traditionally attacked using Skorokhod embedding.
- CAUTION: If (1) has a solution for a positive g, the solution is not necessarily positive:

  Consider X being Brownian motion and $g(x) = x^2$. Then, $f(x) = x^2 - T$.

One dimensional diffusions

- In order to make the discussion precise let's suppose X is a one-dimensional diffusion on a natural scale on (l, r) with a given boundary behaviour, i.e. $A = \frac{1}{2}\sigma^2 \frac{d^2}{dx^2}$ where σ is a nonnegative measurable function. The associated *speed measure*, m, on (l, r) has density $2/\sigma^2(x)dx$ on the interior, I, of the domain.

One dimensional diffusions

- In order to make the discussion precise let's suppose X is a one-dimensional diffusion on a natural scale on (l, r) with a given boundary behaviour, i.e. $A = \frac{1}{2}\sigma^2 \frac{d^2}{dx^2}$ where σ is a nonnegative measurable function. The associated *speed measure*, m, on (l, r) has density $2/\sigma^2(x)dx$ on the interior, I, of the domain.
- Then, it is well-known that A is symmetric with respect to m.
- That is, if f and h are in $\mathcal{D}(A) \cap L^2(I, m) \cap S$, where S is the class of functions satisfying the boundary conditions stipulated by the behaviour of the diffusion at the boundary, then

$$(Af, h) = (f, Ah),$$

where $(f, h) := \int_I f(x)h(x)m(dx)$ is the inner product on the Hilbert space $\mathcal{H} = L^2(I, m)$.

The spectrum of A

- Consider the eigenvalue problem

$$A\phi = -\lambda\phi, \qquad \phi \in S.$$

  The values of λ for which the above equation has a nontrivial solution make up the spectrum of A.
- If none of the boundaries are natural it was shown by Joanne Elliott (1954) that the spectrum of A is discrete.
- Moreover, there exists a function $p(t,x,y)$ symmetric in x and y and continuous in all variables such that for any continuous f vanishing at the boundaries of I we have

$$P_t f(x) = \int_I p(t,x,y)m(dy).$$

- The function $p(t,x,y)$ has the representation

$$p(t,x,y) = \sum_{n=0}^{\infty} e^{-\lambda_n t}\phi_n(x)\phi_n(y),$$

where) $\leq \lambda_0 \leq \lambda_1 \leq \dots \lambda_n \leq \dots$ are the eigenvalues increasing to ∞ and ϕ_ns are corresponding eigenfunctions. Moreover, (ϕ_n) is an orthonormal sequence, i.e $\|\phi_n\| = 1$ and $(\phi_n, \phi_m) = 0$ for $n \neq m$.

- The function $p(t, x, y)$ has the representation

$$p(t, x, y) = \sum_{n=0}^{\infty} e^{-\lambda_n t} \phi_n(x)\phi_n(y),$$

where) $\leq \lambda_0 \leq \lambda_1 \leq \ldots \lambda_n \leq \ldots$ are the eigenvalues increasing to ∞ and ϕ_ns are corresponding eigenfunctions. Moreover, (ϕ_n) is an orthonormal sequence, i.e $\|\phi_n\| = 1$ and $(\phi_n, \phi_m) = 0$ for $n \neq m$.

- The spectral expansion of the transition density in presence of natural boundaries is obtained by H. P. McKean. In this case the spectrum is not necessarily discrete and the eigenfunctions are not necessarily square integrable.

Introducing the inverse operator

- To make the presentation simpler let's redefine P_t so that $P_t f(x) = \int_I p(t,x,y) f(y) m(dy)$. (P_t) is still a kernel but it is sub-Markovian, i.e. $P_t 1 < 1$ if one of the boundaries are absorbing.

Introducing the inverse operator

- To make the presentation simpler let's redefine P_t so that $P_t f(x) = \int_I p(t, x, y) f(y) m(dy)$. (P_t) is still a kernel but it is sub-Markovian, i.e. $P_t 1 < 1$ if one of the boundaries are absorbing.
- To ease the exposition we will assume that there are no natural boundaries but all the results will continue to hold when a natural boundary is present after appropriate modifications.

Introducing the inverse operator

- To make the presentation simpler let's redefine P_t so that $P_t f(x) = \int_I p(t,x,y) f(y) m(dy)$. (P_t) is still a kernel but it is sub-Markovian, i.e. $P_t 1 < 1$ if one of the boundaries are absorbing.
- To ease the exposition we will assume that there are no natural boundaries but all the results will continue to hold when a natural boundary is present after appropriate modifications.
- Note that if $f \in \mathcal{H}$ and vanishes at the absorbing boundaries, then

$$P_t f = \sum_{n=0}^{\infty} e^{-\lambda_n t} (\phi_n, f) \phi_n.$$

- Thus, if $g = P_t f$, we can recover f from g via

$$f = \sum_{n=0}^{\infty} e^{\lambda_n t} (\phi_n, g) \phi_n.$$

- The above gives us a recipe to solve our inverse problem.
- In order to determine the domain of P_t^{-1} first note that $g \in S \cap \mathcal{H} \cap \mathcal{D}(A^\infty)$ if $g = P_t f$ for $t > 0$ whenever $f \in \mathcal{H}$.

- The above gives us a recipe to solve our inverse problem.
- In order to determine the domain of P_t^{-1} first note that $g \in S \cap \mathcal{H} \cap \mathcal{D}(A^\infty)$ if $g = P_t f$ for $t > 0$ whenever $f \in \mathcal{H}$.
- Let $\mathcal{D}(P_t^{-1})$ be the class of functions, g, in $S \cap \mathcal{H} \cap \mathcal{D}(A^\infty)$ such that
$$\sum_{n=0}^{\infty} e^{2\lambda_n t}(g, \phi_n)^2 < \infty.$$
Then, we may define
$$P_t^{-1} : g \in \mathcal{D}(P_t^{-1}) \mapsto \sum_{n=0}^{\infty} e^{\lambda_n t}(g, \phi_n)\phi_n.$$
- It is worth to note that the sum in the above definition converges absolutely and uniformly on the compact squares of $I \times I$.

- Note, however, that since $\lambda_n \to \infty$, $e^{\lambda_n t} \to \infty$, indicating that P_t^{-1} is an unbounded operator.

- Note, however, that since $\lambda_n \to \infty$, $e^{\lambda_n t} \to \infty$, indicating that P_t^{-1} is an unbounded operator.
- Unboundedness of the operator in particular implies that one cannot have a solution for every g and the domain of the operator P_T^{-1} is dense.
- The last observation is both bad and good news: Even if we cannot have a solution to the original problem, we can get arbitrarily close to a solution.

- Note, however, that since $\lambda_n \to \infty$, $e^{\lambda_n t} \to \infty$, indicating that P_t^{-1} is an unbounded operator.
- Unboundedness of the operator in particular implies that one cannot have a solution for every g and the domain of the operator P_T^{-1} is dense.
- The last observation is both bad and good news: Even if we cannot have a solution to the original problem, we can get arbitrarily close to a solution.
- **Remark:** If A were bounded, i.e. if X were a Markov chain, then P_t^{-1} would have the full domain being a bounded operator.

An example

The simplest example of functions in $D(P_t^{-1})$ is $g = p(u, z, \cdot)$ for $u > t$ and $z \in I$. Indeed, since

$$g = \sum_{n=0}^{\infty} e^{-\lambda_n u} \phi_n(z) \phi_n,$$

An example

The simplest example of functions in $D(P_t^{-1})$ is $g = p(u, z, \cdot)$ for $u > t$ and $z \in I$. Indeed, since

$$g = \sum_{n=0}^{\infty} e^{-\lambda_n u} \phi_n(z) \phi_n,$$

$$\sum_{n=0}^{\infty} e^{2\lambda_n t} (g, \phi_n)^2 =$$

An example

The simplest example of functions in $D(P_t^{-1})$ is $g = p(u, z, \cdot)$ for $u > t$ and $z \in I$. Indeed, since

$$g = \sum_{n=0}^{\infty} e^{-\lambda_n u} \phi_n(z) \phi_n,$$

$$\sum_{n=0}^{\infty} e^{2\lambda_n t} (g, \phi_n)^2 = \sum_{n=0}^{\infty} e^{-2\lambda_n (u-t)} \phi_n^2(z)$$

An example

The simplest example of functions in $D(P_t^{-1})$ is $g = p(u, z, \cdot)$ for $u > t$ and $z \in I$. Indeed, since

$$g = \sum_{n=0}^{\infty} e^{-\lambda_n u} \phi_n(z) \phi_n,$$

$$\sum_{n=0}^{\infty} e^{2\lambda_n t} (g, \phi_n)^2 = \sum_{n=0}^{\infty} e^{-2\lambda_n (u-t)} \phi_n^2(z) = p(2(u-t), z, z) < \infty.$$

Moreover,

$$P_t^{-1} g(x) = \sum_{n=0}^{\infty} e^{-\lambda_n (u-t)} \phi_n(z) \phi_n(x) = p(u-t, z, x),$$

An example

The simplest example of functions in $D(P_t^{-1})$ is $g = p(u, z, \cdot)$ for $u > t$ and $z \in I$. Indeed, since

$$g = \sum_{n=0}^{\infty} e^{-\lambda_n u} \phi_n(z) \phi_n,$$

$$\sum_{n=0}^{\infty} e^{2\lambda_n t} (g, \phi_n)^2 = \sum_{n=0}^{\infty} e^{-2\lambda_n (u-t)} \phi_n^2(z) = p(2(u-t), z, z) < \infty.$$

Moreover,

$$P_t^{-1} g(x) = \sum_{n=0}^{\infty} e^{-\lambda_n (u-t)} \phi_n(z) \phi_n(x) = p(u-t, z, x),$$

as one would also guess from the semi-group relation

$$p(u, z, x) = \int_I p(t, x, y) p(u-t, z, y) m(dy)$$

in view of the symmetry of p.

How to determine whether $g \in \mathcal{D}(P_t^{-1})$

- It turns out that there is a 'simple' criterion that will provide the answer to our question.
- Define the operator $\mathcal{J}^\alpha$ for $\alpha > 0$ by

$$\mathcal{J}^\alpha g(t,x) := \int_0^\infty J_0(2\sqrt{ts}) e^{-\alpha s} P_s g(x) ds,$$

  where J_0 is the Bessel function of the first kind of order 0. If X is transient α can be taken equal 0.
- One can show easily that $t \mapsto (\mathcal{J}^\alpha g(t,\cdot), g)$ is convex and decreasing to 0 as $t \to \infty$.

- Coming back to our original question of whether g belongs to $\mathcal{D}(P_T^{-1})$, the answer lies in the tail behaviour of $(\mathcal{J}^\alpha g(t,\cdot), g)$.
- In fact, it is fairly easy to show that $g \in \mathcal{D}(P_T^{-1})$ if and only if

$$\int_0^\infty I_0(2\sqrt{2Tt}) \int_0^\infty J_0(2\sqrt{ts}) e^{-\alpha s}(P_s g, g) ds dt < \infty,$$

where I_0 is the modified Bessel function of the first kind of order 0.
- We also have the inversion formula

$$P_T^{-1} g = e^{-\alpha T} \int_0^\infty I_0(2\sqrt{Tt}) \int_0^\infty J_0(2\sqrt{ts}) e^{-\alpha s} P_s g ds dt. \quad (3)$$

Alternative construction using a Picard iteration

- For $\lambda \geq 0$ consider the λ-potential operator:

$$U^{\lambda} := \int_0^{\infty} e^{-\lambda t} P_t dt.$$

- Then, $\mathcal{J}^{\alpha} g$ is the unique solution of the following Cauchy problem:

$$\begin{aligned} \frac{d}{dt} j(t, \cdot) &= -U^{\alpha} j(t, \cdot) \\ j(0, \cdot) &= U^{\alpha} g. \end{aligned}$$

Alternative construction using a Picard iteration

- For $\lambda \geq 0$ consider the λ-potential operator:

$$U^\lambda := \int_0^\infty e^{-\lambda t} P_t dt.$$

- Then, $\mathcal{J}^\alpha g$ is the unique solution of the following Cauchy problem:

$$\begin{aligned} \frac{d}{dt} j(t,\cdot) &= -U^\alpha j(t,\cdot) \\ j(0,\cdot) &= U^\alpha g. \end{aligned}$$

- Consequently, if we set

$$\begin{aligned} j_0(s,\cdot) &= U^\alpha g, \\ j_{n+1}(s,\cdot) &= U^\alpha g - \int_0^t U^\alpha j_n(r,\cdot)dr, \end{aligned}$$

$(j_n(\cdot,\cdot))_{n\geq 0}$ converges in $L^2([0,t] \times \mathbf{E}, ds \times m)$ to $(\mathcal{J}_s^\alpha g)_{s\in[0,t]}$ if $t < \alpha$, where ds denotes the Lebesgue measure on $[0,t]$.

Regularisation

The most common method in practice is the *Tikhonov regularisation*, which in our set up corresponds to the solution of

$$P_t f + \gamma f = g, \qquad \gamma > 0,\ g \in L^2(m).$$

The above is a special case of the following:

Theorem 1

Suppose that $\phi : \mathbb{R}_+ \mapsto \mathbb{R}_+$ *is continuous with* $\liminf_{x \to \infty} \phi(x) > 0$ *and* $\sup_{x \geq 0} e^{-tx}\phi(x) < \infty$*. Then,*

$$(1-\gamma)P_t f + \gamma \phi(-A) f = g, \qquad \gamma \in (0,1), \tag{4}$$

has a unique solution for any $g \in L^2(m)$ *and* $t > 0$*. Moreover, the solution has the property that*

$$(1-\gamma) f = \arg\min_{h \in L^2(m)} \|P_t h - g\|^2 + \frac{\gamma}{1-\gamma}(P_t \phi(-A) h, h). \tag{5}$$

- The assumption that $\liminf_{x\to\infty} \phi(x) > 0$ cannot be dispensed easily if (4) is to have a solution for any given $g \in L^2(m)$. To wit take $\phi(x) = e^{-tx}$. Then (4) becomes $P_t f = g$, which does not have a solution in general.
- For each $\gamma \in (0,1)$ denote by f_γ the solution of (4). Assume further that $g \in \mathcal{D}(P_t^{-1})$. Then

$$\lim_{\gamma\to 0} \|f_\gamma - P_t^{-1} g\| = 0.$$

- Although looking abstract the above theorem furnishes us with a plethora of concrete examples for regularising the inverse problem. Before giving some concrete examples let us consider the following corollary.

Mixing with jump processes

Suppose that K is a bounded positive operator such that $K = \psi(-A)$ for some bounded continuous function $\psi : \mathbb{R}_+ \mapsto [0, 1]$ and $\gamma \in (0, 1)$ is arbitrary. In an enlargement of the probability space there exists a Markov process Y such that

$$Y = \xi X + (1 - \xi)J,$$

where ξ is a Bernoulli random variable with $\mathbb{P}(\xi = 1) = \gamma$, J is a jump Markov process with generator

$$Bf = Kf - f, \qquad f \in L^2(m),$$

and ξ, J and X are mutually independent. Moreover, $\mathcal{D}(Q_t^{-1}) = L^2(m)$, where (Q_t) is the semigroup associated to Y.

Examples

- Fix $T^* > 0$ and let $K = P_{T^*}$. Then, J is a Markov jump process that remains constant between the jumps of a Poisson process with unit parameter and moves between the points of **E** according to the transition function P_{T^*} or is sent to the cemetery state with probability $1 - \int_{\mathbf{E}} p(T^*, x, y) m(dy)$.
- In particular, when X is a Brownian motion, the process Y is a Brownian motion with probability γ and a compound Poisson process with normally distributed jumps with probability $1 - \gamma$.

Examples

- Fix $T^* > 0$ and let $K = P_{T^*}$. Then, J is a Markov jump process that remains constant between the jumps of a Poisson process with unit parameter and moves between the points of **E** according to the transition function P_{T^*} or is sent to the cemetery state with probability $1 - \int_{\mathbf{E}} p(T^*, x, y)m(dy)$.
- In particular, when X is a Brownian motion, the process Y is a Brownian motion with probability γ and a compound Poisson process with normally distributed jumps with probability $1 - \gamma$.
- One can also consider $K = \lambda U^\lambda$ for some $\lambda > 0$.

What is missing?

A comparison result:

- Recall that Falkner (1983) has shown (under duality assumption and another mild condition) for a general transient Markov process, X, with potential operator U that if $U\mu \leq U\nu$ for measures μ and ν, then one can find a stopping time τ such that X_τ has law μ if ν is the distribution of X_0.

What is missing?

A comparison result:

- Recall that Falkner (1983) has shown (under duality assumption and another mild condition) for a general transient Markov process, X, with potential operator U that if $U\mu \leq U\nu$ for measures μ and ν, then one can find a stopping time τ such that X_τ has law μ if ν is the distribution of X_0.
- A comparison result of this nature would be really useful.

What is missing?

A comparison result:

- Recall that Falkner (1983) has shown (under duality assumption and another mild condition) for a general transient Markov process, X, with potential operator U that if $U\mu \leq U\nu$ for measures μ and ν, then one can find a stopping time τ such that X_τ has law μ if ν is the distribution of X_0.
- A comparison result of this nature would be really useful.
- Note that in order for $g \in D(P_T^{-1})$ it is necessary that $Ug \leq Uh$ for some $h \in D(P_T^{-1})$.

What is missing?

A comparison result:

- Recall that Falkner (1983) has shown (under duality assumption and another mild condition) for a general transient Markov process, X, with potential operator U that if $U\mu \leq U\nu$ for measures μ and ν, then one can find a stopping time τ such that X_τ has law μ if ν is the distribution of X_0.
- A comparison result of this nature would be really useful.
- Note that in order for $g \in D(P_T^{-1})$ it is necessary that $Ug \leq Uh$ for some $h \in D(P_T^{-1})$.
- However, this necessary condition is not sufficient: Let g be the distribution of X_τ , where $\tau = \inf\{t \geq T : |X_t| > a\}$ and X is a standard Brownian motion killed as soon as it exits a finite interval. g cannot be in $D(P_T^{-1})$ since it has a point mass.