Solution to Homework #3, 36-754

25 February 2006

Exercise 10.1

I need one last revision of the definition of a Markov operator: a linear operator on L_1 satisfying the following conditions.

1. If $f \geq 0$ (μ-a.e.), then $Kf \geq 0$ (μ-a.e.).
2. If $f \leq M$ (μ-a.e.), then $Kf \leq M$ (μ-a.e.).
3. $K\mathbf{1}_\Xi = \mathbf{1}_\Xi$.
4. If $f_n \downarrow 0$, then $Kf_n \downarrow 0$.

First, we show that kernels induce operators, then that operators induce kernels.

Given a kernel κ, we define the operator $Kf(x) \equiv \int \kappa(x, dy)f(y)$.

1. Clearly, if $f \geq 0$ a.e., then $\int \kappa(x, dy)f(y) \geq 0$.
2. If $f \leq M$ a.e., then $M - f(x) \geq 0$ a.e., and $\int \kappa(x, dy)(M - f(y)) = \int \kappa(x, dy)M - \int \kappa(x, dy)f(y) = M - Kf(x) \geq 0$ a.e., so $M \geq Kf(x)$ a.e.
3. $\int \kappa(x, dy)\mathbf{1}_\Xi(y) = 1$ for all x, so $K\mathbf{1}_\Xi(x) = \mathbf{1}_\Xi(x)$.
4. If $f_n(x) \downarrow 0$ pointwise, then, for each x, $\int \kappa(x, dy)f_n(x) \to 0$ by monotone convergence, so $Kf_n(x) \to 0$.

Now for the converse: given an operator K, define $\kappa(x, B) = K\mathbf{1}_B(x)$. For fixed x, we need this to be a probability measure. For every $B \in \mathcal{X}$, $\mathbf{1}_B \in L_1$, so $K\mathbf{1}_B(x)$ is, for fixed x, a set-function defined over the whole of the σ-algebra in question. Since $1_{\mathbf{B}}(x) \geq 0$, we know that $K\mathbf{1}_B(x) \geq 0$, and this is a non-negative set-function. Next, we check that it's finitely additive: If A and B are disjoint sets, $\mathbf{1}_{A\cup B}(x) = \mathbf{1}_A(x) + \mathbf{1}_B(x)$. Hence, by the linearity of K, $K\mathbf{1}_{A\cup B} = K\mathbf{1}_A + K\mathbf{1}_B$, which by induction extends to any finite collection of disjoint sets. We notice that if $A_n \downarrow \emptyset$, $\mathbf{1}_{A_n}(x) \downarrow 0$ pointwise, so $K\mathbf{1}_{A_n}(x) \to 0$ (for fixed x). But, by proposition 32 (ch. 3, p. 14), this implies that $K\mathbf{1}_A$ is a measure. Finally, $K\mathbf{1}_\Xi(x) = 1$, for all x, so the measure is a probability measure. We also need $\kappa(x, B)$ to be a measurable function of x, for every fixed set B. Since K takes bounded, measurable functions to bounded, measurable functions, it follows that $K1_B(x)$ is measurable for every x, as required.

What's that business about preserving the L_1 norm of positive functions?

Let me apologize for leading you down various garden paths; I managed to get myself confused about the difference between evolution operators acting to the right, moving functions forward in time, and acting to the left, moving distributions forward in time.

What I asked you to show was that, if $f \in L_1(\mu)$ and $f \geq 0$ μ-a.e., then $\|f\| = \|Kf\|$, where $Kf(x) = \int \kappa(x, dy)f(y)$ for some transition kernel κ. Consider $f(x) = \mathbf{1}_B(x)$, for some $B \in \mathcal{X}$. Clearly, this is non-negative and of finite L_1-norm, in fact $\|\mathbf{1}_B\| = \mu(B)$. Then

$$\begin{aligned}
\|Kf\| &= \int \mu(dx)|Kf(x)| && (1)\\
&= \int \mu(dx)\left|\int \kappa(x, dy)f(y)\right| && (2)\\
&= \int \mu(dx)\left|\int \kappa(x, dy)\mathbf{1}_B(y)\right| && (3)\\
&= \int \mu(dx)\,|\kappa(x, B)| && (4)\\
&= \int \mu(dx)\kappa(x, B) && (5)\\
&= \mu\kappa(B) && (6)
\end{aligned}$$

according to our definition (101, ch. 9, p. 50) of products of kernels. Hence the desired property holds only if μ is invariant under κ. Conversely, if μ is invariant under κ, then this holds for all indicator functions of measurable sets, and so (by linearity) for all simple functions, and so (by the usual arguments) for all integrable, i.e., L_1, functions. Thus, the invariance of μ under κ is a necessary and sufficient condition for $\|Kf\|_\mu = \|f\|_\mu$ when $f \geq 0$ μ-a.e.

A number of people tried the following ingenious argument. $Kf(x) = \mathbf{E}[f(X_1)|X_0 = x]$. If $f \geq 0$, then $\|f\| = \mathbf{E}[f(X)]$. Now use smoothing:

$$\begin{aligned}
\|Kf\| &= \mathbf{E}[Kf(X)]\\
&= \mathbf{E}[\mathbf{E}[f(X_1)|X_0 = X]]\\
&= \mathbf{E}[f(X_1)]\\
&= \|f\|
\end{aligned}$$

Unfortnately, for the last step to be valid for all f, it must be the case that X_1 and X_0 have the same distribution, i.e., that μ is invariant under the action of the Markov process.

There is a version of what I was trying to assert which is correct, regardless of whether or not μ is invariant, but we need to have the operators act to the *left* rather than to the *right*. Given a transition kernel κ and a probability measure

ν, we have defined $\nu\kappa$ as a new probability measure,

$$\nu\kappa(B) = \int \nu(dy)\kappa(y, B)$$

Now suppose that ν is absolutely continuous with respect to some other measure μ. We can write

$$\nu\kappa(B) = \int \mu(dy)\frac{d\nu}{d\mu}\kappa(y, B)$$

by the Radon-Nikodym theorem, which makes it clear that $\nu\kappa$ is also absolutely continuous with respect to μ. The derivative $\frac{d\nu}{d\mu}$ is a function $f \in L_1(\mu)$, such that $f(x) \geq 0$ a.e. μ, and $\|f\| = 1$. Conversely, any such function f is the Radon-Nikodym derivative of a measure ν absolutely continuous with respect to μ. Hence we may write $f\kappa$ to denote the new density obtained by the action *to the left* of the transition kernel κ. But any function with finite $L_1(\mu)$ norm is the Radon-Nikodym derivative of *some* signed measure absolutely continuous with respect to μ. So, we can extend $f\kappa$ to work for arbitrary functions in $L_1(\mu)$:

$$f\kappa(B) = \int \mu(dy)f(y)\kappa(y, B)$$

It follows that this is a linear operator: for any $a, b \in \mathbb{R}$ and all $f, g \in L_1(\mu)$,

$$\begin{aligned}
(af + bg)\kappa(B) &= \int \mu(dy)(af(y) + bg(y))\kappa(y, B) \\
&= a\int \mu(dy)f(y)\kappa(y, B) + b\int \mu(dy)g(y)\kappa(y, B) \\
&= af\kappa(B) + bg\kappa(B)
\end{aligned}$$

We have just seen above that, if $f \geq 0$ and $\|f\| = 1$, then $\|f\kappa\| = 1$, because, in that case, $f\kappa$ is another probability density. But if $f \geq 0$ and $\|f\| = c$, then $\|\frac{1}{c}f\| = 1$, so by linearity $\|f\kappa\| = \|f\|$ if $f \geq 0$.

I apologize for my inability to tell my right from my left.

Exercise 10.2

First, we go from the Markov property (and the associated transition kernels) to the existence of operators for the conditional expectations. For any integrable function f of a random variable X and σ-algebra $\mathcal{F}$, if there is a regular conditional probability $\mathbb{P}(X|\mathcal{F})$, then $\mathbf{E}[f(X)|\mathcal{F}] = \int f\mathbb{P}(X|\mathcal{F})$. For a Markov process, the conditional distribution of X_s given $\mathcal{F}_t^X$ is always regular, and given by $\mu_{t,s}$. Hence

$$\mathbf{E}\left[f(X_s)|\mathcal{F}_t^X\right] = \mathbf{E}\left[f(X_s)|X_t\right] = \int \mu_{t,s}(X_t, dy)f(y) = K_{t,s}f(X_t)$$

as desired.

To go the other way, it's enough to show that $\mathbf{E}\left[f(X_s)|\mathcal{F}_t^X\right] = K_{t,s}f(X_t)$ implies the Markov property, that $\mathbb{P}\left(X_s \in B|\mathcal{F}_t^X\right) = \mathbb{P}\left(X_s \in B|X_t\right)$ for every set $B \in \mathcal{X}$. Consider then $f(x) = \mathbf{1}_B(x)$. Then $\mathbf{E}\left[\mathbf{1}_B(X_s)|\mathcal{F}_t^X\right] = \mathbb{P}\left(X_s \in B|\mathcal{F}_t^X\right) = K_{t,s}\mathbf{1}_B(X_t)$. That last is a function of B and X_t. Since it is clearly $\sigma(X_t)$-measurable, it must be a version of the conditional probability $\mathbb{P}\left(X_s \in B|X_t\right)$. Hence the Markov property holds.

(This is simpler than what I had in mind when I wrote the problem, and more especially the hint. My proof was to show that $K_{t,s}$ induces all the right finite-dimensional distributions. First I showed that the one-dimensional distributions evolved correctly, i.e., according to Eq. 9.3 and 9.4, and then I showed that if all the n-dimensional distributions followed Eq. 9.5, then the $n+1$-dimensional distributions did too. This was unnecessarily complex.)