Notes on Iterated Expectations

Stephen Morris

February 2002

1. Introduction

Consider the following sequence of numbers. Individual 1's expectation of random variable X; individual 2's expectation of individual 1's expectation of X; 1's expectation of 2's expectation of 1's expectation of X; and so on. Samet (1998a) showed that if the individuals have a common prior on a finite state space, then this seqeuence will converge to the same number (which will be equal to the expectation of X contingent on public information). Furthermore, such convergence will occur for all random variables X only if their beliefs are derivable from a common prior.

Samet's argument exploits the fact that such iterated expectations can be represented by a Markov process. In this note, we show how similar techniques can be used to provide a partial characterization of a weaker convergence property. Consider the following sequence of numbers. Individual 1's expectation of random variable X; 1's expectation of 2's expectation of 1's expectation of X; 1's expectation of 2's expectation of 1's expectation of 2's expectation of 1's expectation of X; and so on. Thus we are looking at those iterated expectations that begin and end with individual 1. Using the Markov characterization of iterated expectations, one can show that such convergence is equivalent to showing that, from any initial conditions, a certain Markov process converges to a certain long run distribution.

In this note, I show by example that convergence may fail if there are finite state, but individuals' priors have disjoint supports. I show that *any sequence* may result on an infinite state space but with disjoint supports. A final example shows that convergence may fail with an infinite state space and a common prior, as long as the prior is improper.

Various sufficient conditions for convergence based on well-known properties of Markov processes are reported. On a finite state space, convergence occurs if there is a common support. On an infinite state space, a sufficient condition is that for every event measurable with respect to individual 1's information, there exists $\varepsilon > 0$ such that either that event or its complement have the property that 1's expectation of 2's probability of that event is at least ε. This condition is always true with continuous densities on a continuum state space that are uniformly bounded below.

2. The Main Question: Iterated Expectations from One Player's Perspective

Fix a measure space $(\Omega, \mathcal{F})$. Individual i has a prior measure P_i and a sub σ-field $\mathcal{F}_i$. For any measurable random variable X, we write

$$E_1(X) = \mathbf{E}^{P_i}\left[X | \mathcal{F}_i\right].$$

Let

$$f(X) = E_1(E_2(X))$$

We are interested in identifying conditions under which $f^k(E_1(X))$ converges for all random variables X. I.e., for any X, there exists an $\mathcal{F}_1$-measurable random variable X_1^* such that $f^k(E_1(X))[\omega] \to X_1^*[\omega]$ for all ω.

3. Three Counterexamples

Three counterexamples show a failure of the sequence $\left\{f^k(E_1(X))\right\}_{k=0}^{\infty}$ to converge under different assumptions.

3.1. Example 1

The first example shows that we may fail to get convergence with finite types but disjoint supports in the individuals' prior beliefs.

$$\begin{aligned}
\Omega &= \{\omega_1, \omega_2, \omega_3, \omega_4\} \\
\mathcal{Q}_1 &= (\{\omega_1, \omega_2\}\{\omega_3, \omega_4\}) \\
\mathcal{Q}_2 &= (\{\omega_1, \omega_3\}\{\omega_2, \omega_4\}) \\
P_1 &= \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right) \\
P_2 &= \left(0, \frac{1}{2}, \frac{1}{2}, 0\right)
\end{aligned}$$

Here, individual i's σ-field is generated by partition $\mathcal{Q}_i$. Let

$$X = (x_1, x_1, x_2, x_2).$$

Observe that since X is $\mathcal{Q}_1$-measurable, we have

$$E_1(X) = X.$$

Now

$$E_2(X) = (x_2, x_1, x_2, x_1)$$

and

$$E_1(E_2(X)) = (x_2, x_2, x_1, x_1).$$

Thus

$$\begin{aligned}(E_1 E_2)^{2k} E_1(x) &= (x_1, x_1, x_2, x_2) \\ (E_1 E_2)^{2k+1} E_1(x) &= (x_2, x_2, x_1, x_1)\end{aligned}$$

for all $k = 0, 1, \ldots.$

3.2. Example 2

The second example shows how the sequence $\left\{f^k(E_1(X))\right\}_{k=0}^{\infty}$ may take any value in $\mathbb{R}^{\infty}$, with countably infinite types and disjoint supports.

$$\begin{aligned}\Omega &= \{\omega_1, \omega_2, \ldots\} \\ \mathcal{Q}_1 &= (\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4, \omega_5\}, \ldots\ldots) \\ \mathcal{Q}_2 &= (\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \ldots) \\ P_1 &= \left(\varepsilon, 0, \varepsilon(1-\varepsilon), 0, \varepsilon(1-\varepsilon)^2, 0, \varepsilon(1-\varepsilon)^3, \ldots\right) \\ P_2 &= \left(0, \varepsilon, 0, \varepsilon(1-\varepsilon), 0, \varepsilon(1-\varepsilon)^2, 0, \varepsilon(1-\varepsilon)^3, ..\right)\end{aligned}$$

Now consider any random variable

$$X = \{x_1, x_2, x_2, x_3, x_3, \ldots\ldots\}$$

Observe that since X is $\mathcal{Q}_1$-measurable, we have

$$E_1(X) = X.$$

Now

$$E_2(X) = (x_2, x_2, x_3, x_3, x_4, \ldots.)$$

and

$$E_1(E_2(X)) = (x_2, x_3, x_3, x_4, x_4, \ldots.).$$

Thus

$$(E_1E_2)^k E_1(x) = (x_{k+1}, x_{k+2}, x_{k+2}, x_{k+3}, x_{k+3}, ...)$$

for all $k = 0, 1, ...$

Thus

$$\left\{f^k(E_1(X))\right\}_{k=0}^{\infty}[\omega_1] = \{x_1, x_2, x_3, ...\}.$$

3.3. Example 3

A final example shows how we may fail to get convergence with uncountably infinite types and a common (but improper) prior.

$$\begin{aligned}
\Omega &= \{\omega_1, \omega_2, ...\} \\
\mathcal{Q}_1 &= (\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4, \omega_5\}, ...) \\
\mathcal{Q}_2 &= (\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, ...) \\
P_1 &= P_2 = \left(1, \frac{\alpha}{1-\alpha}, \left(\frac{\alpha}{1-\alpha}\right)^2, \left(\frac{\alpha}{1-\alpha}\right)^3, ...\right)
\end{aligned}$$

where $\alpha > \frac{1}{2}$. Now consider any random variable

$$X = \{x_1, x_2, x_2, x_3, x_3, ...\}$$

Observe that since X is $\mathcal{Q}_1$-measurable, we have

$$E_1(X) = X.$$

Now

$$E_2(X) = (\alpha x_2 + (1-\alpha)x_1, \alpha x_2 + (1-\alpha)x_1, \alpha x_3 + (1-\alpha)x_2, \alpha x_3 + (1-\alpha)x_2, ...)$$

and

$$E_1(E_2(X)) = \left(\begin{array}{l} \alpha x_2 + (1-\alpha)x_1, \\ \alpha^2 x_3 + 2\alpha(1-\alpha)x_2 + (1-\alpha)^2 x_1, \\ \alpha^2 x_3 + 2\alpha(1-\alpha)x_2 + (1-\alpha)^2 x_1, ... \end{array}\right).$$

Now consider an increasing sequence of integers $1 = n_1 < n_2 < ...$ and let

$$X(\omega) = \begin{cases} 0, \text{ if } n_{2k+1} \le \omega < n_{2k} \\ 1, \text{ if } n_{2k} \le \omega < n_{2k+1} \end{cases}$$

for all $k = 0, 1, ...$

For any fixed ε, one can choose n_1 and k_1 such that

$$f^{k_1}(E_1(X))[\omega_1] < \varepsilon$$

(independent of $n_2, n_3, ...$). One can then choose n_2 and k_2 such that

$$f^{k_2}(E_1(X))[\omega_1] > 1 - \varepsilon$$

(independent of $n_3, n_4, ...$). One can choose n_3 and k_3 such that

$$f^{k}(E_1(X))[\omega_1] < \varepsilon$$

(independent of $n_4, n_5, ...$). And so on.

4. Discrete State Spaces

Let Ω be a countable set. Each individual's information can be represented by a partition $\mathcal{Q}_i$. Each individual has a prior P_i. We allow for the possibility that P_i is improper, but require that each element in i's partition has finite probability (so conditional probabilities will be well-defined). We can label the elements of i's partition by the natural numbers, so that $\mathcal{Q}_i = (Q_{i1}, Q_{i2}, ...)$. Now we can write $p_i(n,m)$ for the ex ante probability that i assigns to 1 observing Q_{1n} and 2 observing Q_{2m}. Thus

$$p_i(n,m) = \sum_{\omega \in Q_{1n} \cap Q_{2m}} P_i(\omega).$$

Write $p_i(m|n)$ for the condition probability that i assigns to j observing Q_{jm} if he has observe Q_{in}.

Let N_i be the number of elements of i's partition (N_i may equal ∞). Let M^i be the $N_i \times N_j$ matrix whose (n,m)th element is $p_i(m|n)$. Clearly, this is a Markov matrix.

Let

$$M = M^1 M^2$$

be the $N_1 \times N_1$ product of these two matrices. Notice that M is also a Markov process, with (n, n')th element equal to

$$\sum_{m=1}^{N_2} p_1(m|n)\, p_2(n'|m).$$

Now observe that any $\mathcal{Q}_1$-measurable random variable X can be naturally represented as a N_1-vector x (with the interpretation that the x_n is the realization of X on the event Q_{1n}).

But now if

$$x = E_1(X),$$

we have that

$$f^k(E_1(X)) = M^k x.$$

Proposition 4.1. *$f^k(E_1(X))$ converges everywhere for all X if and only if the sequence of Markov matrices M^k converges.*

The Markov process corresponding to example 1 deterministically jumps from one type to the other. Thus there are cycles in iterated expectations. The Markov process corresponding to example 2 deterministically moves up states monotonically (and never converges or cycles). The Markov process corresponding to example 3 is a biased random walk.

Corollary 4.2. *If either player has a finite number of types and there is a common support, then $f^k(E_1(X))$ converges everywhere for all X.*

The corollary follows from properties of Markov matrices.

Observe that example 1 fails the sufficient condition because there is not common support, while examples 2 and 3 fail because both individuals have an infinite number of types.

5. The General Case

In general, we define a transition function $Q : \Omega \times \mathcal{F}_1 \to [0,1]$ by

$$Q(\omega, Z) = \int_{\omega' \in F} P_2[Z|\omega'] \, P_1[\omega'|\,\omega] \, d\omega'.$$

Thus $Q(\omega, \cdot)$ is a measure on $(\Omega, \mathcal{F}_1)$ for all $\omega \in \Omega$ and $Q(\cdot, Z)$ is $\mathcal{F}_1$-measurable for all $Z \in \mathcal{F}_1$.

Transition process Q describes a Markov process on $(\Omega, \mathcal{F}_1)$. We write Q^n for the Markov process generated by n transitions. Now we have the following condition

Condition M. There exists $\varepsilon > 0$ and integer $N \geq 1$ such that for any $Z \in \mathcal{F}_1$, either $Q^N(\omega, Z) \geq \varepsilon$ for all ω; or $Q^N(\omega, Z^c) \geq \varepsilon$ for all ω.

Stokey and Lucas (1989) show that this condition allows a contraction mapping argument that can be used to show the existence of a unique invariant measure with exponential convergence to that measure. Thus we have:

Proposition 5.1. *If condition M is satisfied, then $f^k(E_1(X))$ converges everywhere for all X.*

An easy sufficient condition for condition M is:

Condition M_1. There exists $\varepsilon > 0$ such that for any $Z \in \mathcal{F}_1$, either $Q(\omega, Z) \geq \varepsilon$ for all ω; or $Q(\omega, Z^c) \geq \varepsilon$ for all ω.

We can interpret this condition in the special case of continuous densities. Let $\Omega = [0,1]^2$ with $\omega = (\omega_1, \omega_2)$. Let $\mathcal{F}_i$ be generated by i observing ω_i only and let f_i be individual i's prior density. Now

$$Q(\omega, Z) = \int_{\omega_2 \in [0.1]} \int_{\omega_1' \in [0,1]} f_2[\omega_1' | \omega_2] f_1[\omega_2 | \omega_1] d\omega_1' d\omega_2.$$

A sufficient condition for this to satisfy M_1 is that $f_2[\omega_1 | \omega_2]$ is uniformly bounded below. In this case, there exists $\varepsilon > 0$ such that

$$Q(\omega, Z) = \int_{\omega_2 \in [0.1]} \int_{\omega_1' \in Z} f_2[\omega_1' | \omega_2] f_1[\omega_2 | \omega_1] d\omega_1' d\omega_2 \geq \int_{\omega_1' \in Z} \varepsilon d\omega_1'.$$

We can no doubt weaken condition M and still get our convergence. An open question is whether convergence is necessary with *proper* common priors.

6. Samet's Question Revisited: The Common Prior and Order Independent Convergence

We look at iterated expectations from 1's perspective, and asked when $f^k(E_1(X))$ converges for all random variables X. I.e., for any X, there exists an $\mathcal{F}_1$-measurable random variable X_1^* such that $f^k(E_1(X))[\omega] \to X_1^*[\omega]$ for all ω. Of course, we could also define

$$g(X) = E_2(E_1(X))$$

and ask when, for any X, there exists an $\mathcal{F}_2$-measurable random variable X_2^* such that $g^k(E_2(X))[\omega] \to X_2^*[\omega]$ for all ω. If we have convergence from each individual's perspective, we can then ask if the limits are the same for the two players, i.e., $X_1^* = X_2^*$. This is the problem that Samet solved for the finite case.

Another characterization of the common prior for the finite case is that two individuals have a common prior, for all X, $E_1(X) \geq 0$ and $E_2(X) \geq 0$ implies $E_1(X) = 0$ and $E_2(X) = 0$ (Morris (1994)). This result extends to the infinite case only under some compactness assumptions (Feinberg (2001), Samet (1998b), Ng (1998)). Presumably, these infinite characterizations and the results in this paper could be used to provide an infinite state analogue of Samet (1998a).

References

[1] Feinberg, Y. (2001). "Characterizing Common Priors in the Form of Posteriors," *Journal of Economic Theory* 91, 127-179.

[2] Morris, S. (1994). "Trade with Heterogeneous Priors and Asymmetric Information," *Econometrica* 62, 1327-1347.

[3] Nehring, K. (2001). "Common Priors under Incomplete Information: A Unification," *Economic Theory* 18, 535-553.

[4] Ng, M. (1998). "A Simple Proof for the Existence of Common Priors."

[5] Samet, D. (1998a). "Iterated Expectations and Common Priors," *Games and Economic Behavior* 24, 131-141.

[6] Samet, D. (1998b). "Common Priors and the Separation of Convex Sets," *Games and Economic Behavior* 24, 172-174.

[7] Stokey, N. and R. Lucas (1989). *Recursive Methods in Economic Dynamics.* Harvard University Press.