DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

MERGING OF FORECASTS IN MARKOV MODELS

In Ho Lee
University of Southampton

SOCIAL SCIENCE WORKING PAPER 951

January 1996

Merging of Forecasts in Markov Models

In Ho Lee

Abstract

Blackwell and Dubins (1962) and Kalai and Lehrer (1994) showed that absolute continuity is necessary and sufficient for merging of opinions. This paper suggests the concept of merging of forecasts which is a modification of merging of opinions in Markov models where the underlying state of nature may change over time.

We define the merging of forecasts as the conditional probabilities of the future state given the past observations of signals drawn conditional on the state get close to each other for different agents; it allows for the event that agents agree on the future evolution of the states even if they have not agreed in the distant past. For an ergodic Markov chain, any forecasts merge. In particular, we can dispense with the absolute continuity for merging of forecasts.

Merging of Forecasts in Markov Models

In Ho Lee*

1 Introduction

Consider a group of investors who forecast the prospect of a firm as more information about the firm becomes available over time. Will they agree about the prospect as more information becomes available? How different priors can they start with and yet arrive at an agreement? The present paper attempts to answer these questions in a setting where the prospect of the firm changes over time in a Markov fashion.

Since Blackwell and Dubins (1962) showed that absolute continuity is sufficient for merging of opinions, Kalai and Lehrer (1994) showed that absolute continuity is also necessary. Absolute continuity requires that the priors of investors be compatible in the sense that they agree on which states are possible as the initial state.

The paper proposes that the definition of merging of opinions be modified in a model where the underlying state of nature changes over time in Markov fashion to allow for the event that agents agree on the future evolution of the system even if they do not agree on the past. In particular I propose the concept of merging of forecasts which represents the event where conditional probabilities of the future state given the past observations of signals drawn conditional on the state get close to each other for different agents.

Equipped with the definition, a stronger result on the merging is obtained. If all states form a single ergodic class, merging of forecasts is obtained without absolute continuity on the initial priors; to arrive at an agreement investors may at the beginning totally disagree in the sense that different investors put probability 1 on different states. For instance investor A believes that the firm is in the good state with probability 1 while investor B believes that the firm is in the bad state with probability 1. Yet after many time periods, the two investors agree on the probability of the current state being the good state and the bad state if they observe the same sequence of information.

*I acknowledge the financial support from ESRC Research Fellowship. Remaining errors are mine.

2 Merging of Forecasts

Let $\{S_t, t = 0, 1, \dots\}$ be a Markov chain on $S = \{s^1, \dots, s^K\}$ with the transition law P. Let $(\Sigma, \mathcal{E})$ be the measurable space generated by the Markov chain. We denote the measure on $(\Sigma, \mathcal{E})$ consistent with P by μ.

An agent observes signals Z_t drawn conditionally *i.i.d* given the current state S_t, but not directly S_t. In particular the agent knows that $Z_t \sim f(\cdot|S_t)$ where $f(\cdot|S_t)$ is the conditional distribution. Denoting the measurable space generated by $\{Z_t, t = 1, \dots\}$ as $(\Omega, \mathcal{F})$, we denote the measure on $(\Omega, \mathcal{F})$ by ν. We assume that the signal Z_t is not perfectly informative of S_t : i) supp $f(\cdot|s) = F$, a measurable subset of $\Re$, for all $s \in S$ and ii) $0 < f(\cdot : s) < \infty$ for all $Z \in F$. A filtration on $(\Omega, \mathcal{F})$ is a sequence of the set of history $\mathcal{F}_t = \{Z_1, \dots Z_t\}$ such that i) for all t, $\mathcal{F}_t \in \mathcal{F}$, and $\mathcal{F}_{t+1}$ refines $\mathcal{F}_t$ and ii) $\lim_{t\to\infty} \mathcal{F}_t = \mathcal{F}$. The product measure $\mu \times \nu$ on the product space $(\Sigma \times \Omega, \mathcal{E} \otimes \mathcal{F})$ is well defined in a natural way.

In period t, the agent wants to forecast the future conditional on $\mathcal{F}_t$. Initially the agent is endowed with the prior on S_0, $\widetilde{\mu}_0 = (\widetilde{\mu}_0^1, \dots, \widetilde{\mu}_0^K)$ while the true prior is denoted as $\mu_0 = (\mu_0^1, \dots, \mu_0^K)$. We denote the probability measure consistent with $\widetilde{\mu}_0$ as $\widetilde{\mu}$ and thus $\widetilde{\mu} \times \nu$ for the probability measure that the agent works with.

We denote the projection of Σ for $t \geq n$ by Σ^n and similarly the projection of Ω, Ω^n. The sets Σ^n and Ω^n accompanied by the $\sigma-$algebras $\mathcal{E}^n$ and $\mathcal{F}^n$contain only the future evolution of the system from period n and on. We define that $\mu \times \nu$ is absolutely continuous with respect to $\widetilde{\mu} \times \nu$ from n, denoted $\mu \times \nu \ll^n \widetilde{\mu} \times \nu$, if for all $A \in \mathcal{E}^n \otimes \mathcal{F}^n$, there is $\delta > 0$ such that $\mu \times \nu(A) \leq \delta \cdot (\widetilde{\mu} \times \nu(A))$.

The agent's conditional distribution of an event $A \in \mathcal{E} \otimes \mathcal{F}$ given $\mathcal{F}_t$ is denoted as $\mu \times \nu\{A|\mathcal{F}_t\}$. The a*gent's forecast* $\widetilde{\mu} \times \nu$ *merges to* $\mu \times \nu$ if for every $\varepsilon > 0$, and almost every $\omega \in \mathcal{E} \otimes \mathcal{F}$, there is $t(\varepsilon, \omega)$ satisfying $|\widetilde{\mu} \times \nu\{S_{t+1}|\mathcal{F}_t\} - \mu \times \nu\{S_{t+1}|\mathcal{F}_t\}| < \varepsilon$ for every $t \geq t(\varepsilon, \omega)$.

Theorem 1 *Suppose* $\{S_t, t = 0, 1, \dots\}$ *admits a single ergodic class. Then the agent's forecast* $\widetilde{\mu}$ *merges to* μ.

The proof is accomplished through a few steps.

Proposition 1 *If* $\{S_t, t = 0, 1, \dots\}$ *admits a single ergodic class, then there is* n *such that* $\mu \times \nu$ *is absolutely continuous with respect to* $\widetilde{\mu} \times \nu$ *from* n.

Proof: First we show that if $\{S_t, t = 0, 1, \dots\}$ admits a single ergodic class, then $\widetilde{\mu}_t^k > 0$ implies that $\widetilde{\mu}_{t+1}^k > 0$. Suppose not, namely $\widetilde{\mu}_t^k > 0$ implies that $\widetilde{\mu}_{t+1}^k = 0$. Since Z_t is not perfectly informative, the observation of Z_t cannot make $\widetilde{\mu}_{t+1}^k = 0$. Therefore $p_{kk} = 0$ where p_{kk} is the probability from the state k to itself and $\widetilde{\mu}_t^j = 0$ for all $j \in J$ where J is the set of immediate predecessors in the Markov chain, that is $p_{jk} > 0$ for all $j \in J$. The same argument is applied to all states in J in period t, that is, $p_{j'j} = 0$ for all $j, j' \in J$ and $\widetilde{\mu}^i_{t-1} = 0$ for all $i \in I$ where I is the set of immediate predecessors of the states in J. This argument can be applied only a finite times before the process returns to the state k since $\{S_t\}$ is a recurrent Markov chain and thus there is $\eta \leq K$ such that $p^\eta_{kk} > 0$ and $p^\iota_{kk} = 0$ for all $\iota < \eta$. A contradiction since $\{S_t\}$ is aperiodic.

Second there is $n \in N$ such that $\widetilde{\mu}^k_n > 0$ for all k since all states are recurrent.

It remains to show that there is n such that for all $A \in \mathcal{E}^n \otimes \mathcal{F}^n$, there is $\delta > 0$ such that $\mu \times \nu(A) \leq \delta \cdot (\widetilde{\mu} \times \nu(A))$. Consider n that $\widetilde{\mu}^k_n > 0$ for all k. Then for $A \in \mathcal{E}^n \otimes \mathcal{F}^n$,

$$\begin{aligned}
\mu \times \nu(A) &= \sum_{k=1}^{K} \mu^k_n \Pr\{A|S_n = s^k\} \\
&\leq M \sum_{k=1}^{K} \Pr\{A|S_n = s^k\} \\
&= \frac{M}{m}(m \sum_{k=1}^{K} \Pr\{A|S_n = s^k\}) \\
&\leq \frac{M}{m}(\sum_{k=1}^{K} \widetilde{\mu}^k_n \Pr\{A|S_n = s^k\}) \\
&= \frac{M}{m}(\widetilde{\mu} \times \nu(A))
\end{aligned}$$

where $M = \max_k\{\mu^k_n\}$ and $m = \min_k\{\widetilde{\mu}^k_n\}$. Taking $\delta = \frac{M}{m}$ completes the proof. ∎

Proposition 2 *If $\mu \times \nu$ is absolutely continuous with respect to $\widetilde{\mu} \times \nu$, then $\widetilde{\mu} \times \nu$ merges to $\mu \times \nu$.*

For proof, see Blackwell and Dubins (1962) and Kalai and Lehrer (1994).

Proof: (Theorem 1) From Proposition 1, $\mu \times \nu \ll^n \widetilde{\mu} \times \nu$. Applying Proposition 2 to the projection of the process from period n and on, namely $(\Sigma^n \times \Omega^n, \mathcal{E}^n \otimes \mathcal{F}^n)$, completes the proof since $\{S_{t+1}\} \in \mathcal{E}^n \otimes \mathcal{F}^n$, $t \geq n$. ∎

3 Merging of Forecasts and Merging of Opinions

Consider a problem whether a jury can agree on the verdict of a suspect. Individual jurors are required to have *unbiased* opinions before the trial in the sense that they are prepared to allow either conviction or acquittal based on the evidence being fed to them in the court. Otherwise a juror's eventual opinion may not depend on the evidence provided in the court but it may be entirely determined by the prior. Hence the verdict may not be consistent with the evidence. To avoid this, jurors are required to have an unbiased opinion initially. Using the terminology of Kalai and Lehrer (1993), their probability measures must contain a grain of truth.

However, in the problem of investors, they do not have to agree on the past event in order to agree on the future prospect of the firm. As long as they are aware that the firm's prospect may change over time, observing a long string of common data enables agents to agree on the future prospect. In this circumstance the past is relevant only to the extent that it helps agents forecasts the future. In particular a disagreement as to the past does not prevent an agreement as to the future.

The difference between merging of forecasts and merging of opinions lie in that the former explicitly restricts the event to be considered to the future ones; merging of opinions requires that probability measures on all events including the past as well as the future get close to each other and thus it is a stronger condition.

The restriction of events to the future in the definition of merging of forecasts yields a case where merging of forecasts holds true while merging of opinions does not. In particular in the problem of investors, merging of opinions does not hold true without the absolute continuity on the initial priors although merging of forecasts holds true regardless of the initial priors. It happens only when the future can be different from the past.

Indeed casting the statistical inference problem due to Blackwell and Dubins (1962) into a Markov chain sheds lights on the issue. We can regard the statistical inference problem as the one where each possible value of unknown parameter constitutes an absorbing state of a Markov chain since the unknown parameter is fixed once and for all. In this circumstance, the prediction of the future evolution of signals which are drawn conditional on the parameter will get close to each other only if the agents start with priors which are absolutely continuous with each other. Hence employing the concept of merging of forecasts instead of merging of opinions does not yield any different result.

We can regard an ergodic Markov chain as a model with built-in grain-of-truth. In this class of models, absolute continuity among probability measures exist inherently in the system since one cannot permanently exclude a realization of any state in the stochastic process; even if one does initially, there is a strictly positive probability that the process visits any state. The observation is important since the failure of merging of opinions results from the fact that Bayesian updating *per se* cannot put a positive probability measure on an event whose prior is zero.

The role of aperiodicity in the proof of proposition 1 suggests an interesting example. For instance, consider a Markov chain with alternating two states so that it has the periodicity of 2. In this case two agents with the opposite degenerate priors may never agree; although they realize that there is a positive probability of visiting each state, they assign probability 1 to either of the state in any particular period, hence leaving no room for Bayesian updating to work on the belief conditional on the observation of the signal draws.

Recently increasing attention has been paid to dynamic models where the state changes in Markov fashion. For models in this class, our result has a positive implication. Although the issue of merging of forecasts or merging of opinions is rarely addressed in those models explicitly, it can be a very important underlying issue if divergent opinions change result drastically which is typical in models with heterogeneity among agents. Furthermore the ever-changing nature of the state could introduce additional complexity in the model. Our positive result implies that in the class of models frequently used in economics, an even stronger result can be obtained. In particular, absolute continuity is not required if the states constitute a single ergodic class.

References

[1] Blackwell, D., and L. Dubins, Merging of Opinions with Increasing Information, *Annals of Mathematical Statistics,* Vol. 38, (1962), 882-886.

[2] Doob, J. L., *Stochastic Processes*, New York, John Wiley and Sons, (1953).

[3] Kalai, E., and E. Lehrer, Rational Learning Leads to Nash Equilibrium, *Econometrica*, Vol. 61, (1993), 1019-1045.

[4] ------ and --------, Weak and Strong Merging of opinions, *Journal of Mathematical Economics,* Vol. 23, (1994), 73-86.