Robustness and Information Processing

Kenneth Kasa

Review of Economic Dynamics 9(2006)

February 16, 2007

Motivation

- The paper tries to relate two apparently different literatures: Robust Control (RC) and Rational Inattention (RI).

Motivation

- The paper tries to relate two apparently different literatures: Robust Control (RC) and Rational Inattention (RI).
- Why might we want to do that?

Motivation

- The paper tries to relate two apparently different literatures: Robust Control (RC) and Rational Inattention (RI).
- Why might we want to do that?

  Lucas' admonition: *'Beware of theorists bearing free parameters'*.

Motivation

- The paper tries to relate two apparently different literatures: Robust Control (RC) and Rational Inattention (RI).
- Why might we want to do that?

  Lucas' admonition: *'Beware of theorists bearing free parameters'*.

  $\Rightarrow$ Information-theoretic approach to the calibration of robustness (θ).

Logic

- Starting point: RC and RI are based on alternative interpretation of measurement error:
  - RC: Measurement error viewed as subject to misspecification chosen by the evil agent.
  - RI: Measurement error as unavoidable and rational due to the presence of limited information processing capacity of the individual.
- A Kalman filtering problem (in continuous time) is analyzed under alternative assumptions about model uncertainty and information processing.
- Notice the symmetry: evil agent in a robust filter *maximizes* forecast errors subject to an *upper bound*, whereas an agent with finite Shannon capacity *minimizes* forecast errors subject to a *lower* bound.
- The idea of the paper: to parameterize the maximum degree of model uncertainty using the minimum degree of Shannon capacity.

Standard Continuous-Time Kalman Filtering Basics

- Probability space $(\Omega, \mathcal{F}, P)$ on which filtration $\mathcal{F}_t = \sigma\left(X_0, W^1(s), W^2(s); 0 \leq s \leq t\right)$ is defined.
- Consider the pair of stochastic differential equations

$$\begin{aligned} dX &= -aXdt + \sigma dW^1 \\ dY &= Xdt + dW^2 \end{aligned}$$

- The agent's filtering problem can be expressed as

$$\inf_{\widehat{X}_t \in \mathcal{X}_t} \|X_t - \widehat{X}_t\|^2 = \inf_{\widehat{X}_t \in \mathcal{X}_t} \int_\Omega \left(X_t - \widehat{X}_t\right)^2 dP$$

- The solution of this problem can be written in the following recursion

$$d\widehat{X} = -a\widehat{X}dt + K(t)\left[dY - \widehat{X}dt\right]$$

- Kalman gain satisfies the Riccati equation $\dot{K} = -2aK - K^2 + \sigma^2$.
- As $t \to \infty$, $K(t)$ converges to $\overline{K} = -a + \sqrt{a^2 + \sigma^2}$.

Incorporating Uncertainty

- Let $Q, P \in \mathcal{M}(\Omega)$ and $Q \ll P$.
- Acknowledging model misspecification

$$
\begin{aligned}
dX &= -\left(aX - \sigma\xi_1(t)\right)dt + \sigma d\widetilde{W}^1 \\
dY &= \left(X + \xi_2(t)\right)dt + d\widetilde{W}^2.
\end{aligned}
$$

where $\widetilde{W}^i(t) = W^i(t) - \int_0^t \xi_i(s)\,ds$ for $i = 1, 2$.

- Relative entropy bound

$$
h_\infty(Q\|P) \equiv \lim_{T\to\infty} \sup \frac{1}{2T} E^Q \int_0^T \left(|\xi_1(s)|^2 + |\xi_2(s)|^2\right) ds < \mathcal{U}
$$

Robust Filter cont'd

A robust filter can be characterized as the Nash equilibrium of the following zero-sum game

$$V_t = \inf_{\{\widehat{X}_s\}} \sup_{Q} \left\{ \lim_{T\to\infty} \sup \frac{1}{T} E^Q \int_0^T \left| X_s - \widehat{X}_s \right|^2 ds - \theta h_\infty (Q \| P) \right\}$$

Proposition If $\theta \geq \sigma^2 / \left(\sigma^2 + a^2\right)$ there is a unique solution to the robust

filtering problem given by

$$\begin{aligned} d\widehat{X} &= -a\widehat{X}dt + K_r(t)\left[dY - \widehat{X}dt\right] \\ \dot{K}_r &= -2aK_r - \left(1 - \theta^{-1}\right) K_r^2 + \sigma^2 \end{aligned}$$

where the robust Kalman filter gain, $K_r(t)$, converges to the following solution of the algebraic Riccati equation, $\dot{K}_r = 0$

$$\overline{K}_r = \frac{-a + \sqrt{a^2 + \left(1 - \theta^{-1}\right)\sigma^2}}{1 - \theta^{-1}}$$

Rational Inattention: Main Ideas

Rational Inattention is based on the idea that economic agents cannot pay attention to *all* available information but can decide *which* information to pay attention to.

Rational Inattention: Main Ideas

Rational Inattention is based on the idea that economic agents cannot pay attention to *all* available information but can decide *which* information to pay attention to.

Definition Let Y and X be two random processes on $[0, T]$ with joint probability measure $\mu_{X,Y}$ and marginal probability measures μ_Y and μ_X. If $\mu_{X,Y} \ll \mu_X \times \mu_Y$ then the *mutual information*, $\mathcal{I}_T(X, Y)$, between Y and X is

$$\mathcal{I}_T(X, Y) = \int \ln\left(\frac{d\mu_{X,Y}}{d\left[\mu_X \times \mu_Y\right]}\right) d\mu_{X,Y}$$

Rational Inattention: Main Ideas

Rational Inattention is based on the idea that economic agents cannot pay attention to *all* available information but can decide *which* information to pay attention to.

Definition Let Y and X be two random processes on $[0, T]$ with joint probability measure $\mu_{X,Y}$ and marginal probability measures μ_Y and μ_X. If $\mu_{X,Y} \ll \mu_X \times \mu_Y$ then the *mutual information*, $\mathcal{I}_T(X, Y)$, between Y and X is

$$\mathcal{I}_T(X, Y) = \int \ln\left(\frac{d\mu_{X,Y}}{d[\mu_X \times \mu_Y]}\right) d\mu_{X,Y}$$

The idea is that agents can reduce the conditional variance of the state by exercising some monitoring effort. This effort is limited, however, by the constraint that the flow of information must not exceed the channel capacity, κ.

Rational Inattention: Capacity Constrained Filtering

Proposition The mutual information of the transmission channel is given by

$$\mathcal{I}_T(X, Y) = \frac{1}{2} E \int_0^T \left(X_s - \widehat{X}_s \right)^2 ds$$

where $\widehat{X}_s = E(X_s \mid Y_u; 0 \leq u \leq s)$ and is given by the (regular) Kalman filter.

Rational Inattention: Capacity Constrained Filtering

Proposition The mutual information of the transmission channel is given by

$$\mathcal{I}_T(X, Y) = \frac{1}{2} E \int_0^T \left(X_s - \widehat{X}_s \right)^2 ds$$

where $\widehat{X}_s = E(X_s \mid Y_u; 0 \leq u \leq s)$ and is given by the (regular) Kalman filter.

Corollary The steady-state rate of information conveyed by Y about X is given by

$$\lim_{T \to \infty} \frac{1}{T} \mathcal{I}_T(X, Y) = \frac{1}{2} \overline{K} = \frac{1}{2} \left(-a + \sqrt{a^2 + \sigma^2} \right)$$

RC vs. RI: Observational Equivalence

Proposition If $\kappa < \frac{1}{2}\left(-a + \sqrt{a^2 + \sigma^2}\right)$ then Rational Inattention and Robust Filtering are observationally equivalent, in the sense that a higher filter gain can either be interpreted as an increased preference for robustness or an increased ability to process information.

RC vs. RI: Observational Equivalence

Proposition If $\kappa < \frac{1}{2}\left(-a + \sqrt{a^2 + \sigma^2}\right)$ then Rational Inattention and Robust Filtering are observationally equivalent, in the sense that a higher filter gain can either be interpreted as an increased preference for robustness or an increased ability to process information.

$\Rightarrow$ The robust filter gain can be reinterpreted as giving rise to a 'demand' for information processing as a function of θ

$$\kappa^d(\theta) = \frac{-a + \sqrt{a^2 + (1 - \theta^{-1})\,\sigma^2}}{2(1 - \theta^{-1})}$$

Mapping κ and θ: "Methodology"

- Start from the likelihood ratio test: $H_0 : dY = X_{0t}dt + dW_t$ vs. $H_1 : dY = X_{1t}dt + dW_t$.
- Define the moment-generating function of $l(T) \equiv \ln \Lambda(T)$ as

$$M_i(s) = E\{\exp[sl(T)|H_i]\} = E[\Lambda(T)^s|H_i]$$

- Then a standard Chernoff bound calculation yields $P(error|H_i) \leq M_i(s)$.
- Using some "tricks" from probability theory delivers

$$M_T(s) = \exp\left[\frac{s^2 - s}{2}\int_0^T P(t)\,dt\right]$$

- Noting that $s = 1/2$ is the minimizing choice of s and letting $T \to \infty$ gives the following detection error probability $P(error) \leq \exp\left[-\frac{1}{8}P_\infty T\right]$.
- Relate P_∞ to the capacity of an information transmission channel.

Mapping κ and θ for Anderson, Hansen and Sargent (2003)

The reference model has the following representation

$$\begin{aligned} dx_{1t} &= (\mu_0 + \mu_1 x_{2t})\, dt + \sigma_1 dW_{1t} \\ dx_{2t} &= -\mu_2 x_{2t} dt + \sigma_2 dW_{2t} \end{aligned}$$

Assume that x_{2t} is an unobserved hidden state.
Calculate detection error probabilities

$$d \leq \frac{1}{2} \exp\left\{ -\frac{1}{8} |g|^2 \, T \right\}$$

where $g = -\delta\theta^{-1} [\sigma_1 v_1 + \sigma_2 v_2]$ and where v_1 and v_2 are coefficients on x_{1t} and x_{2t} from the value function.

Mapping κ and θ cont'd

A standard result in signal processing gives us

$$d \leq \frac{1}{2} \exp\left\{ -Ts \left(\kappa - \frac{\mu_1 \sigma_2}{4\sigma_1 \sqrt{1-s}} \right) \right\}$$

Relate channel capacity and robustness as in the previous proposition

$$\theta^{-1} = \frac{\mu_1^2}{\sigma_1^2} \left[1 + \frac{\mu_2}{\kappa} - \frac{\mu_1^2 \sigma_2^2}{4\sigma_1^2 \kappa^2} \right]$$

Mapping κ and θ cont'd

Table 1
Calibration using the AHS approach

d	$\|g\|$	θ^{-1}	θ
0.01	0.396	19.8	0.050
0.05	0.303	15.1	0.066
0.10	0.254	12.7	0.079
0.15	0.219	10.9	0.091
0.20	0.191	9.55	0.105
0.25	0.166	8.30	0.120
0.30	0.143	7.15	0.140
0.40	0.094	4.70	0.213

Table 2
Calibration using an information-theoretic approach

d	κ_{nats}	κ_{bits}	θ^{-1}	θ	$\|g\|$
0.01	0.079	0.114	10.9	0.091	0.218
0.05	0.067	0.097	8.94	0.112	0.179
0.10	0.061	0.088	7.40	0.135	0.148
0.15	0.058	0.084	6.42	0.156	0.128
0.20	0.056	0.081	5.67	0.176	0.113
0.25	0.054	0.078	4.83	0.207	0.097
0.30	0.053	0.076	4.37	0.229	0.087
0.40	0.051	0.073	3.36	0.298	0.067

Mapping κ and θ cont'd

Table 2
Calibration using an information-theoretic approach

d	κ_{nats}	κ_{bits}	θ^{-1}	θ	$\lvert g \rvert$
0.01	0.079	0.114	10.9	0.091	0.218
0.05	0.067	0.097	8.94	0.112	0.179
0.10	0.061	0.088	7.40	0.135	0.148
0.15	0.058	0.084	6.42	0.156	0.128
0.20	0.056	0.081	5.67	0.176	0.113
0.25	0.054	0.078	4.83	0.207	0.097
0.30	0.053	0.076	4.37	0.229	0.087
0.40	0.051	0.073	3.36	0.298	0.067

Table 3
Calibration using the HSW approach

d	$\lvert g \rvert$	θ^{-1}	θ
0.01	0.396	19.5	0.051
0.05	0.303	14.9	0.067
0.10	0.254	12.5	0.080
0.15	0.219	10.8	0.093
0.20	0.191	9.41	0.106
0.25	0.166	8.18	0.122
0.30	0.143	7.04	0.142
0.40	0.094	4.63	0.216