Robust Inattentive Discrete Choice

Lars Peter Hansen (University of Chicago)
Jianjun Miao (Boston University)
Hao Xing (Boston University)

AEA/ASSA Meeting: Expectations and Macro-Finance

January 8, 2022

Illustration

- Invest in one of a finite number of projects each with an uncertain payoff modeled as exposed to a "hidden state"
- Decide how much to learn about the projects in advance of making the investment
- There are "information costs" associated with the learning
- Uncertain prior distribution over the hidden states

Setup

- $x \in X$ is a state realization (X is a finite set)
- μ is a prior distribution over states in X
- $s \in S$ is a signal realization (S is a finite set)
- $d(s \mid x)$ is the information structure, the vector signal probabilities conditioned on a state x.
- $a \in A$ is a potential action
- $\sigma(s) = a$ is a decision rule that assigns an action to each signal realization
- $u(x, a)$ is a utility function over states and actions

In rational inattention models, d becomes an object of choice subject to information costs along with σ.

Mutual information

- joint distribution $d(s \mid x)\mu(x)$
- two marginals $\mu(x)$ over states and $\bar{d}(s) = \sum_x d(s \mid x)\mu(x)$ over signals
- mutual information is the KL divergence of the joint relative to the two marginals:

$$\sum_x \sum_s d\mu \left[\log(d\mu) - \log\left(\bar{d}\mu\right)\right]$$

Two equivalent representations

- Entropy measures:

$$\mathcal{H}(s) = -\sum_{s} \bar{d}(s) \log \bar{d}(s)$$

$$\mathcal{H}(s \mid x) = -\sum_{x} \sum_{s} d(s \mid x) \log d(s \mid x) \mu(x)$$

  $\mathcal{H}(x)$ and $\mathcal{H}(x \mid s)$ are defined analogously.

- Mutual information:

$$\mathcal{I}(d, \mu) = \mathcal{H}(x) - \mathcal{H}(x \mid s) = \mathcal{H}(s) - \mathcal{H}(s \mid x)$$

- Observations:
  - Measures informational gain in the posterior relative to the prior.
  - convex in d given μ
  - concave in μ given d

Signal-based rational inattention

- Let $\lambda > 0$ denote the shadow cost of information
- Problem:

$$V(\mu) \equiv \max_{d,\sigma} \sum_{x} \sum_{s} d(s \mid x)\mu(x)u\left[x, \sigma(s)\right] \\ - \lambda \mathcal{I}(d, \mu),$$

Observation: solved for a fixed prior $\mu = \hat{\mu}$.

Important references on inattention

Sims (1998, 2003)

Caplin and Dean (2013, 2015)

Matejka and McKay (2015)

Caplin, Dean and Leahy (2019, 2020)

Robust Bayesian approach

To confront uncertainty in a decision problem without information acquisition:

$$\max_{\sigma} \min_{d,\mu} \sum_x \sum_s d(s \mid x)\mu(x)u\left[x, \sigma(s)\right]$$
$$+ \xi \sum_x \sum_s d(s \mid x)\mu(x)\left[\log d(s \mid x) - \log \hat{d}(s \mid x)\right]$$
$$+ \theta \sum_x \mu(x)\left[\log \mu(x) - \log \hat{\mu}(x)\right]$$

where the second term adjusts for model misspecification and the third term adjusts for prior ambiguity relative to a baseline $(\hat{d}, \hat{\mu})$.

Special case of variational preferences: Maccheroni, Marinacci, Rustichini (2006)

Robust signal-based inattention I

- Modify decision theory under uncertainty by allowing the decision maker to choose d subject to a mutual information cost rather than guarding against misspecification
- Modify rational inattention by including a robust choice of μ instead of imposing a baseline $\hat{\mu}$

Robust signal-based inattention II

$$\max_{d,\sigma} \min_{\mu} \sum_x \sum_s d(s \mid x)\mu(x)u\left[x, \sigma(s)\right] \\ - \lambda\mathcal{I}(d, \mu) + \theta \sum_x \mu(x)\left[\log \mu(x) - \log \hat{\mu}(x)\right]$$

While $-\lambda\mathcal{I}(d, \mu)$ is not linear in μ given d, it is convex as is the objective given (σ, d).

A convenient reformulation

- Exchange min μ and max d.
- Fix σ and μ, and construct the partition:

$$S_j = \{s \in S : \sigma(s) = a_j\}$$

- It is optimal to set $d(s \mid x) = 0$ for all but one of the elements of the nonempty S_j. Thus, we may suppose that there is a one-to-one mapping between the signals assigned positive probability and the actions realized by σ. Each signal recommends a course of action.
- Form distribution $p(a \mid x)$ implied by the restricted set of $d(s \mid x)$.
- Re-pose the problem as one with choice-based probabilities.

Robust choice-based inattention

The decision maker selects probability distribution p over actions given states subject to a mutual information cost $\mathcal{I}(p, \mu)$.

$$\max_p \min_\mu \sum_x \sum_a p(a \mid x)\mu(x)u(x, a) - \lambda \mathcal{I}(p, \mu) + \theta \sum_x \mu(x) [\log \mu(x) - \log \hat{\mu}(x)]$$

Concave in p and convex in μ.

Minimax and robust Bayes

When does max – min equal min – max?

- The objective is concave in p given μ and convex in μ given p.

Why do we care?

- Opens the door to a robust Bayesian interpretation: the robust inattention problem is a rational inattention problem for some prior
- Supports convenient characterization and computation

Reversing order of optimization

$$\min_{\mu} \max_{p} \sum_{x} \sum_{a} p(a \mid x)\mu(x)u\left(x, a\right) \\ - \lambda\mathcal{I}(p, \mu) + \theta \sum_{x} \mu(x)\left[\log \mu(x) - \log \hat{\mu}(x)\right]$$

Observe that the inner maximization problem takes μ as given. The robust solution solves a max problem for $\mu = \mu^*$ where μ^* solves the outer minimization problem.

Maximizing by choice of p given μ

- Form the marginal $q^*(a) = \sum_x \mu(x) p^*(a \mid x)$. Then

$$p^*(a \mid x) \propto q^*(a) \exp\left[\frac{u(x,a)}{\lambda}\right].$$

- Additional inequalities accommodate $q^*(a) = 0$
- Form

$$v(x) \doteq \lambda \log \sum_a q^*(a) \exp\left[\frac{u(x,a)}{\lambda}\right]$$

  Then $V(\mu) = \sum_x \mu(x) v(x)$ is the optimized value net of the robustness adjustment.

Exponential tilt towards the high utility states relative to q^*.

Robust choice of μ given p

- Problem:

$$\min_{\mu} \sum_{x} \mu(x) \left[v(x) + \theta \left[\log \mu(x) - \log \hat{\mu}(x)\right]\right]$$

  for v from the max problem.

- Solution:

$$\mu^*(x) \propto \exp\left[-\frac{1}{\theta} v(x)\right]$$

Exponential tilt towards low utility outcomes.

Computation

Propose a convenient algorithm to solve the robust RI problem numerically

- ▹ generalizes the Arimoto (1972) - Blahut (1972) algorithm
- ▹ application of block coordinate descent
  - Iterate back and forth between max and min problems using first-order conditions
  - Include the marginal over actions as part of the iterations

Illustrations

Using illustrations we:

- ▷ explore tradeoff between acting to learn and acting to consume
- ▷ ask when does robustness imitate risk aversion and when does it differ

Motivated by prior analyses of Caplin, Dean and Leahy (2019).

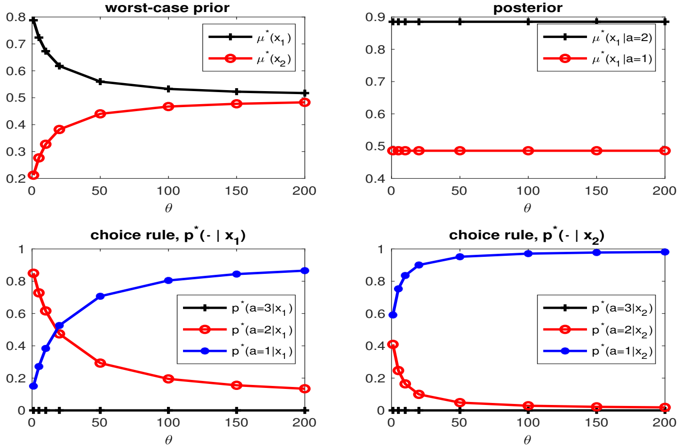
Example

Suppose there are three investments, two states with equal prior probability, and investor has linear utility:

- two are risky and pay off in different states, one higher than the other
- a third is risk-free, constant across states

Action \ State	x_1	x_2
1	0	15
2	6	0
3	5	5

Solutions for alternative concerns about robustness



Observations

- The risk-free investment dominated because of the absence of learning.
- Without robustness concerns, the investor most often prefers the risky investment with the highest possible payoff.
- Robustness considerations push against this dominance as the robust-adjusted prior probabilities assign more weight to the lower payoff state.