How much is information worth? A geometric insight using duality between payoffs and beliefs

Michel De Lara and Olivier Gossner

CERMICS, École des Ponts ParisTech
CNRS, X – Paris Saclay and LSE

Saint Petersburg State University
10 September 2019

I am not sure if my husband is cheating on me What should I do?

A spouse

- can gather information about the current state of Nature:
  has my husband really been to this (mathematical) conference?
  if yes, was his secretary travelling with him?
  is my husband cheating on me?
- makes a decision, taken from a set:
  - stay faithful to her husband ("freeze")
  - stay with her husband and cheat on him ("fight")
  - divorce ("flee")

What is the value of hiring a private detective?
Will valuable information make the spouse change her current choice?

Decision under incomplete information

Investment, insurance, voting, hiring, etc.
virtually all decisions involve incomplete information

How valuable information is depends on

- the agent's available decisions
- the agent's utility function (preferences)
- the agent's prior belief on the state of Nature
- the piece of information

Uniform approach: Blackwell (1951, 1953)

A piece of information α is more informative than β iff
all agents (available decisions, utility, prior) weakly prefer α to β.

Our objective

What is the value of a given piece of information for a given agent?

Outline of the presentation

Outline of the presentation

A Geometric View of the Value of Information

Confident, Undecided, Flexible

Examples: Small Information

Conclusion

An agent acquires information before making a decision

An agent

- observes information about the current state of Nature
- makes a decision, taken from a set

How much information is worth for the agent depends jointly on

- the information provided
- the decision problem (decisions at stake and preferences)

Our objective

Characterize the Value of Information
based on separate conditions on

- the information structure
- the choices available
  (instrumental approach: choice=decision+payoff)

Here is how we frame the problem in mathematical clothes

Prior belief and information received

- A (finite) set K of states of nature, a prior belief $\bar{b} \in \Delta = \Delta(K)$
- An information structure is a random variable (r.v.) $\mathbf{B}$ with values in Δ such that $\mathbb{E}\mathbf{B} = \bar{b}$ (beliefs about beliefs)

Decisions and preferences

Set D of decisions, utility function $u\colon D \times K \to \mathbb{R}$
Actions are payoff vectors $\mathbb{A} = \{u(d, \cdot), d \in D\} \subset \mathbb{R}^K$
We assume $\mathbb{A}$ compact, convex (mixed strategies)

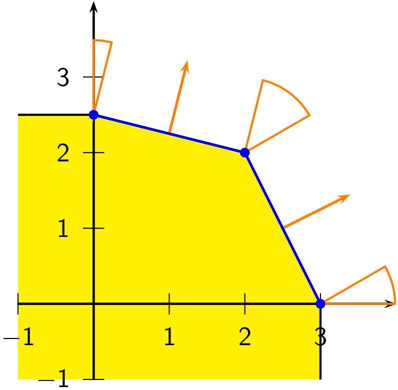
Value of information

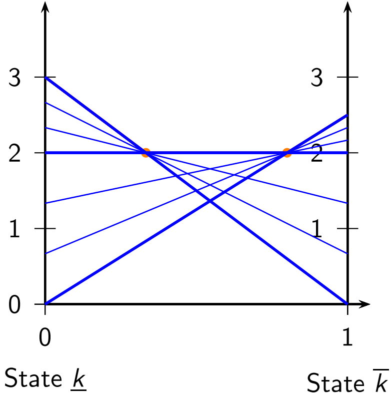
$$V_{\mathbb{A}}(b) = \sup_{a \in \mathbb{A}} \mathbb{E}_b a = \sup_{a \in \mathbb{A}} \langle b \, , a \rangle \ , \ \text{for all belief } b \in \Delta$$

$$\mathrm{VoI}_{\mathbb{A}}(\mathbf{B}) = \mathbb{E} V_{\mathbb{A}}(\mathbf{B}) - V_{\mathbb{A}}(\mathbb{E}\mathbf{B}) \ , \ \text{for all information structure } \mathbf{B}$$

Geometric representation of the *value function* $V_{\mathbb{A}}(b) = \max_{a \in \mathbb{A}} \mathbb{E}_b a = \max_{a \in \mathbb{A}} \langle b\,, a \rangle$

$K = \{\underline{k}, \overline{k}\}$, choices/payoffs are $\{(3,0),(2,2),(0,\frac{5}{2})\}$





Optimal action a as a function of belief b

Belief b is in normal to $\mathbb{A}$ at action a

Varying action a

Geometric formalization using duality

Between actions $\mathbb{A}$ and beliefs Δ, we consider the bilinear pairing

$$\langle b \, , a \rangle = \mathbb{E}_b a \; , \;\; \forall a \in \mathbb{A} \; , \;\; \forall b \in \Delta$$

that is the expected utility of action/payoff a under belief b

Geometric formalization using convex analysis

- The value function $V_{\mathbb{A}}(b) = \max_{a \in \mathbb{A}} \langle b \, , a \rangle$ is
  - the support function of the set $\mathbb{A}$

$$V_{\mathbb{A}} = \sigma_{\mathbb{A}} : \Delta \to \mathbb{R}$$

  - whose sugradient at $b \in \Delta$ is given by

$$\partial V_{\mathbb{A}}(b) = \arg\max_{a \in \mathbb{A}} \langle b \, , a \rangle$$

    the exposed face of $\mathbb{A}$ at b

- The Fenchel conjugate of the value function is
  - the characteristic function of the set $\mathbb{A}$

$$V_{\mathbb{A}}^{\star} = \sigma_{\mathbb{A}}^{\star} = \delta_{\mathbb{A}} : \mathbb{R}^{K} \to \mathbb{R}$$

  - whose sugradient at $a \in \mathbb{A}$ is given by

$$\partial V_{\mathbb{A}}^{\star}(a) = N_{\mathbb{A}}(a) \cap \Delta$$

    where $N_{\mathbb{A}}(a)$ is the normal cone of $\mathbb{A}$ at a

Justifiable actions / Exposed face

Optimal actions

For any belief $b \in \Delta$, let $\mathbb{A}^\star(b)$ be the the set of optimal actions at b (justifiable actions)

$$\mathbb{A}^\star(b) = \{a \in \mathbb{A} \mid V_{\mathbb{A}}(b) = \langle b , a \rangle\} = \arg\max_{a \in \mathbb{A}} \langle b , a \rangle$$

- Optimal actions $\mathbb{A}^\star(b)$ form the exposed face of $\mathbb{A}$ at b, that is, the subgradient of $V_{\mathbb{A}}$ at b

$$\mathbb{A}^\star(b) = \partial V_{\mathbb{A}}(b)$$

- Actions in $\mathbb{A}^\star(b)$ can be justified as they are compatible with belief b

Revealed beliefs / Normal cone

Revealed beliefs

For any action $a \in \mathbb{A}$, let $\Delta^{\star}_{\mathbb{A}}(a)$ be the beliefs revealed by a (justifiable)

$$\Delta^{\star}_{\mathbb{A}}(a) = \{ b \in \Delta \mid \forall a' \in \mathbb{A} \, , \ \langle b \, , a' \rangle \leq \langle b \, , a \rangle \}$$

- Revealed beliefs $\Delta^{\star}_{\mathbb{A}}(a)$ are
  the beliefs in the normal cone of the set $\mathbb{A}$ at action a,
  that is, are related to the subgradient of $V^{\star}_{\mathbb{A}}$ at a by

$$\Delta^{\star}_{\mathbb{A}}(a) = \partial V^{\star}_{\mathbb{A}}(a) \cap \Delta = N_{\mathbb{A}}(a) \cap \Delta$$

- The revealed beliefs $\Delta^{\star}_{\mathbb{A}}(a)$ are compatible with the observed action, hence non refutable

Outline of the presentation

A Geometric View of the Value of Information

Confident, Undecided, Flexible

Examples: Small Information

Conclusion

Information has value if and only if it does impact choices

Confidence set

A belief $b \in \Delta$ is in the confidence set $\Delta^{\mathrm{c}}_{\mathbb{A}}(\bar{b})$ of the prior belief $\bar{b}$ if the optimal actions at $\bar{b}$ are also optimal at b, that is,

$$\Delta^{\mathrm{c}}_{\mathbb{A}}(\bar{b}) = \bigcap_{a \in \mathbb{A}^{\star}(\bar{b})} \Delta^{\star}_{\mathbb{A}}(a)$$

The confidence set $\Delta^{\mathrm{c}}_{\mathbb{A}}(\bar{b})$ is closed, convex and contains $\bar{b}$

Proposition

$$\begin{aligned} \mathrm{VoI}_{\mathbb{A}}(\mathbf{B}) = 0 \quad & \textit{iff} \quad \exists a^{\star} \in \mathbb{A}^{\star}(\bar{b}) \, , \; a^{\star} \in \mathbb{A}^{\star}(\mathbf{B}) \textit{ a.s.} \\ & \textit{iff} \quad \mathbf{B} \in \Delta^{c}_{\mathbb{A}}(\bar{b}) \textit{ a.s.} \end{aligned}$$

This result is aligned with the common wisdom that information is valueless if it does not impact choices

Confident

Theorem: Bounds on the VoI

There exist a positive constant $C_{\mathbb{A}}$ such that,
for every information structure $\mathbf{B}$,

$$C_{\mathbb{A}} \mathbb{E} d(\Delta^{\mathrm{c}}_{\mathbb{A}}(\bar{b}), \mathbf{B}) \geq \mathrm{VoI}_{\mathbb{A}}(\mathbf{B}) \geq \mathrm{VoI}_{\mathbb{A}^{\star}(\bar{b})}(\mathbf{B})$$

where $d(\Delta^{\mathrm{c}}_{\mathbb{A}}(\bar{b}), b') = \inf_{b \in \Delta^{\mathrm{c}}_{\mathbb{A}}(\bar{b})} \|b - b'\|$

Undecided

Proposition

The two following conditions are equivalent

- *There are more than two optimal actions in* $\mathbb{A}^\star(\bar{b})$
- *The value function* $V_{\mathbb{A}}$ *is not differentiable at the prior belief* $\bar{b}$

In that case we say the agent is undecided at $\bar{b}$

Example: indifference in a finite choice set

Bounds on the VoI for the undecided agent

If the agent is undecided at $\bar{b}$, there exist positive constants $C_{\bar{b},\mathbb{A}}$ and $c_{\bar{b},\mathbb{A}}$ such that, for every information structure $\mathbf{B}$,

$$C_{\bar{b},\mathbb{A}}\mathbb{E}\|\mathbf{B}-\bar{b}\| \geq \mathrm{VoI}_{\mathbb{A}}(\mathbf{B}) \geq c_{\bar{b},\mathbb{A}}\mathbb{E}\|\mathbf{B}-\bar{b}\|_{\Sigma^{\mathrm{i}}_{\mathbb{A}}(\bar{b})} \ ,$$

where $\|\cdot\|_{\Sigma^{\mathrm{i}}_{\mathbb{A}}(\bar{b})}$ is a semi-norm with kernel $\left[\mathbb{A}^\star(\bar{b})-\mathbb{A}^\star(\bar{b})\right]^\perp$

The valuable directions of information are the tie-breaking ones

Flexible

Suppose that $\mathbb{A}$ has boundary $\partial\mathbb{A}$ which is a C^2 submanifold of $\mathbb{R}^K$

Proposition

The three following conditions are equivalent:

- *The set-valued mapping $b \mapsto \mathbb{A}^\star(b)$ is a mapping which is a local diffeomorphism at $\bar{b}$*
- *The Hessian of the value function $V_{\mathbb{A}}$ at the prior belief $\bar{b}$ is well defined and is definite positive*
- *The curvature of $\mathbb{A}$ at $\mathbb{A}^\star(\bar{b})$ is positive*

In that case we say the agent is flexible at $\bar{b}$

Examples: portfolio investment, scoring rules.

Theorem: Bounds on the VoI for the flexible agent

If the agent is flexible at $\bar{b}$, there exist positive constants $C_{\bar{b},\mathbb{A}}$ and $c_{\bar{b},\mathbb{A}}$ such that, for every information structure $\mathbf{B}$,

$$C_{\bar{b},\mathbb{A}}\mathbb{E}||\mathbf{B} - \bar{b}||^2 \geq \mathrm{VoI}_{\mathbb{A}}(\mathbf{B}) \geq c_{\bar{b},\mathbb{A}}\mathbb{E}||\mathbf{B} - \bar{b}||^2$$

Confident, Undecided, Flexible

- An agent can be both confident (for certain beliefs) and undecided (in certain directions of information): the value function $V_{\mathbb{A}}$ is not differentiable at belief $\bar{b}$ and displays a flat part (vee shape)
- A flexible agent cannot be confident or undecided: the value function $V_{\mathbb{A}}$ is differentiable at belief $\bar{b}$ and does not display a flat part

Outline of the presentation

A Geometric View of the Value of Information

Confident, Undecided, Flexible

Examples: Small Information

Conclusion

Small information acquisition

- Browsing the web, magasines in a waiting room
- Turning on the radio for a couple of minutes
- Windows shopping
- A quick look at a pile of job applications

Both costs and benefits are relatively low
Can the benefit compensate the cost? (When?)

Notations

Radner-Stiglitz (1984)

Under some technical conditions, the “marginal value” of a little piece of information is null

Letting $(\mathbf{B}^\theta)_{\theta>0}$ be a family of information structures, the marginal value of information is

$$V^+ = \limsup_{\theta\to 0} \frac{1}{\theta}\mathrm{VoI}_{\mathbb{A}}(\mathbf{B}^\theta)$$

Our contribution

Our bounds on the VoI allow to characterise the marginal VoI based on separate conditions on

- the parameterized information structure $(\mathbf{B}^\theta)_{\theta>0}$
- the decision problem at hand $\mathbb{A}$

Setting

In all three following examples,

- we assume binary states of nature $K = \{0, 1\}$
- and we denote by $\bar{b}$ the prior belief on the state being 1

We label as

- confident the case in which $\bar{b}$ lies in the interior of the (closed convex) confidence interval $\Delta^{\mathrm{c}}_{\mathbb{A}}(\bar{b})$
- undecided the case in which the decision maker displays indifference between two actions at $\bar{b}$
- flexible the case in which the optimal action is a smooth function of the belief in a neighborhood of $\bar{b}$

Brownian motion (experimentation, repeated games...)

- Assume the agent observes the realisation of a Brownian motion with variance 1 and drift $k \in \{\underline{k}, \overline{k}\}$ from time 0 to (small) θ

$$d\mathbf{Z}_t = kdt + d\mathbf{W}_t \ , \ \ 0 \leq t \leq \theta$$

- The agent has initially uniform beliefs on the drift $k \in \{\underline{k}, \overline{k}\}$

$$\bar{b} = \frac{1}{2}\delta_{\underline{k}} + \frac{1}{2}\delta_{\overline{k}}$$

- For a small interval of time $\theta > 0$, we have

$$\mathbb{E}\|\mathbf{B}^\theta - \bar{b}\| \sim \sqrt{\theta} \ , \ \ \mathbb{E}\|\mathbf{B}^\theta - \bar{b}\|^2 \sim \theta$$

Marginal value of information

- Confident: $V^+ = 0$
- Undecided: $V^+ = +\infty$
- Flexible: $0 < V^+ < +\infty$

Poisson (multi-armed bandits, strategic experimentation...)

- Assume the agent observes a Poisson process with intensity ρ from time 0 to (small) θ
- The agent has initially uniform beliefs on the intensity $\rho \in \{\underline{\rho}, \overline{\rho}\}$

$$\bar{b} = \frac{1}{2}\delta_{\underline{\rho}} + \frac{1}{2}\delta_{\overline{\rho}}$$

- The observation of a success leads to an a posteriori $b = \frac{\overline{\rho}}{\overline{\rho}+\underline{\rho}}\delta_{\overline{\rho}} + \frac{\underline{\rho}}{\overline{\rho}+\underline{\rho}}\delta_{\underline{\rho}}$ and happens with probability $\sim \theta$
  For a small interval of time $\theta > 0$, we have

$$\mathbb{E}\|\mathbf{B}^{\theta} - \bar{b}\| \sim \theta \ , \ \ \mathbb{E}\|\mathbf{B}^{\theta} - \bar{b}\|^2 \sim \theta$$

Marginal value of information

- Confident:
  - $V^+ = 0$ if b is in the confidence set of $\bar{b}$
  - $0 < V^+ < +\infty$ if b is not in the confidence set of $\bar{b}$
- Undecided: $0 < V^+ < +\infty$
- Flexible: $0 < V^+ < +\infty$

Equally likely signals

- The agent has initially uniform beliefs on $\{\overline{k}, \underline{k}\}$

$$\overline{b} = \frac{1}{2}\delta_{\overline{k}} + \frac{1}{2}\delta_{\underline{k}}$$

- After observing a signal, the equally likely posterior beliefs are

$$(\frac{1}{2} - \theta^{\alpha})\delta_{\overline{k}} + (\frac{1}{2} + \theta^{\alpha})\delta_{\underline{k}}\ ,\ \ (\frac{1}{2} + \theta^{\alpha})\delta_{\overline{k}} + (\frac{1}{2} - \theta^{\alpha})\delta_{\underline{k}}$$

$$\mathbb{E}\|\mathbf{B}^{\theta} - \overline{b}\| \sim \theta^{\alpha}\ ,\ \ \mathbb{E}\|\mathbf{B}^{\theta} - \overline{b}\|^2 \sim \theta^{2\alpha}$$

Marginal value of information

- Confident:
  - $V^+ = 0$
- Undecided:
  - $V^+ = \infty$ if $\alpha < 1$
  - $0 < V^+ < +\infty$ if $\alpha = 1$
  - $V^+ = 0$ is $\alpha > 1$
- Flexible:
  - $V^+ = \infty$ if $\alpha < \frac{1}{2}$
  - $0 < V^+ < +\infty$ if $\alpha = \frac{1}{2}$
  - $V^+ = 0$ is $\alpha > \frac{1}{2}$

Summary of cases

For two elements of x, y of $\mathbb{R}_+ \cup \{\infty\}$, we use the notation $x \simeq y$ if x, y are both 0, both finite and positive (strictly), or both infinite:

$$x \simeq y \iff x, y \in \{(0,0), (\infty,\infty)\} \cup]0,\infty[\times]0,\infty[$$

V^+	Confident	Undecided	Flexible
Poisson	1 (or 0)	1	1
Brownian	0	∞	1
ELS, $\alpha < \frac{1}{2}$	0	∞	∞
ELS, $\alpha = \frac{1}{2}$	0	∞	1
ELS, $\frac{1}{2} < \alpha < 1$	0	∞	0
ELS, $\alpha = 1$	0	1	0
ELS, $\alpha > 1$	0	0	0

Relation with the literature

- Radner, R., and J. Stiglitz (1984): "A nonconcavity in the value of information," in *Bayesian Models of Economic Theory*, ed. by M. Boyer, and R. Kihlstrom, pp. 33–52, Amsterdam. Elsevier. Joint conditions on the parameterized information structure $(\mathbf{B}^{\theta})_{\theta>0}$ and the decision problem at hand $\mathbb{A}$, leading to $V^{+}=0$
- Chade, H., and E. Shlee (2002): "Another look at the Radner-Stiglitz Nonconcavity in the Value of Information," *Journal of Economic Theory*, 107, 421–452. Joint/separate conditions on the parameterized information structure $(\mathbf{B}^{\theta})_{\theta>0}$ and the decision problem at hand $\mathbb{A}$, leading to $V^{+}=0$
- De Lara, M., and L. Gilotte (2007): "A tight sufficient condition for Radner–Stiglitz nonconcavity in the value of information," *Journal of Economic Theory*, 137(1), 696–708. Separate conditions on the parameterized information structure $(\mathbf{B}^{\theta})_{\theta>0}$ and the decision problem at hand $\mathbb{A}$, leading to $V^{+}=0$
- De Lara, M., and O. Gossner
  Separate conditions on the parameterized information structure $(\mathbf{B}^{\theta})_{\theta>0}$ and the decision problem at hand $\mathbb{A}$,
  leading to $V^{+}=\infty$, $0<V^{+}<+\infty$ or $V^{+}=0$

Outline of the presentation

To conclude

The value of information VoI depends on
how strong is the effect of information on choices

- Lowest for a confident decision maker
  (locally flat value function $V_{\mathbb{A}}$)
  The agent is "hard to convince" to change decisions
  The information structure **B** must charge beliefs
  outside the confidence set to "shake" the agent
- Highest in case of an indifference in the choice set
  (kinked value function $V_{\mathbb{A}}$)
  A "small piece" of information can have
  a large influence on the decision
- Mild when the decision problem is smooth and one-to-one
  (curved value function $V_{\mathbb{A}}$)
  In this case, the optimal decision when the belief is $\bar{b}$
  is "almost optimal" (envelope theorem) when the belief is near $\bar{b}$

Open question

- Historically, dual variables have moved
  from geometric (Lagrange) to economic (Kantorovich) flavor
  - Lagrange multipliers of inequality constraints
    are geometric dual variables
  - Kantorovich "resolving multipliers" of constrained primal quantities
    (or "objectively determined estimators")
    are economic dual variables
    (The price of a resource is the sensitivity of the optimal payoff
    with respect to a small increment of the resource)
- In the duality between payoffs/actions and beliefs, what is
  - the equivalent of a production function?
    (is it minus a risk measure?)
  - the "economic" interpretation of beliefs (probability distributions)
    as dual variables of primal payoff/action vectors
    (one payoff per state of the world)?