数理解析研究所講究録 1371 巻 2004 年 110-115

Dempster-Shafer Updating and Additivity of the Core*

Atsushi Kajii
Institute of Economic Research
Kyoto University
kajii@kier.kyoto-u.ac.jp

Takashi Ui
Faculty of Economics
Yokohama National University
oui@ynu.ac.jp

March 29, 2004

Abstract

We discuss an interpretation of the Dempster-Shafer updating rule for belief functions in terms of unanimity games, utilizing the additivity of the core operation. This interpretation leads us to a simple proof for the equivalence of Dempster-Shafer updating rule and the maximum likelihood updating rule.

Key words: capacity; belief function; Dempster-Shafer updating; maximum likelihood; unanimity game; additivity of the core.

1 Notation and basic facts

We use the following notation and definitions.

- Ω is a finite set. Each element $\omega \in \Omega$ is called a state.
- $\mathcal{F} = 2^{\Omega} \setminus \{\emptyset\}$ is the collection of all non-empty subsets of Ω. For each $E \in \mathcal{F}$, we write $\mathcal{F}_E = \{S \in \mathcal{F} : S \cap E \neq \emptyset\}$.
- A function $v : 2^{\Omega} \to \mathbb{R}$ with $v(\emptyset) = 0$ is a *game* and identified with a point in $\mathbb{R}^{\mathcal{F}}$.
- A game $v \in \mathbb{R}^{\mathcal{F}}$ is a *capacity* if it is non-negative ($v(S) \geq 0$ for all $S \in \mathcal{F}$), monotone ($v(S) \leq v(T)$ if $S \subseteq T$), and normalized ($v(\Omega) = 1$).
- A game $v \in \mathbb{R}^{\mathcal{F}}$ is *additive* if $v(S) + v(T) = v(S \cup T)$ for $S, T \in \mathcal{F}$ with $S \cap T = \emptyset$. An additive capacity is called a probability function. The set of all the probability functions is denoted by $\Delta(\Omega)$.

*Kajii acknowledges financial support by MEXT, Grant-in-Aid for 21st Century COE Program. Ui acknowledges financial support by MEXT, Grant-in-Aid for Scientific Research.

- A game $v \in \mathbb{R}^{\mathcal{F}}$ is *convex* (or *supermodular*) if $v(S) + v(T) \leq v(S \cup T) + v(S \cap T)$ for all $S, T \in \mathcal{F}$.

- For $T \in \mathcal{F}$, denote by $u_T \in \mathbb{R}^{\mathcal{F}}$ the *unanimity game* on T: it is a game such that $u_T(S) = 1$ if $T \subseteq S$ and $u_T(S) = 0$ otherwise. A unanimity game is a convex capacity, so is any convex combination of them.

- For $v \in \mathbb{R}^{\mathcal{F}}$, its *conjugate* $v' \in \mathbb{R}^{\mathcal{F}}$ is defined by the rule $v'(S) = v(\Omega) - v(S^c)$ for all $S \in \mathcal{F}$ where $S^c = \Omega \backslash S$. Note that $(v')' = v$ and $(v + w)' = v' + w'$.

- The core of $v \in \mathbb{R}^{\mathcal{F}}$ is

$$C(v) := \{x \in \mathbb{R}^{\Omega} : x(S) \geq v(S) \text{ for all } S \in \mathcal{F} \text{ and } x(\Omega) = v(\Omega)\}$$

where $x(S) = \sum_{\omega \in S} x(\omega)$. If v is a capacity,

$$C(v) = \{p \in \Delta(\Omega) : p(S) \geq v(S) \text{ for all } S \in \mathcal{F}\}.$$

- Note that the core of a unanimity game has a simple structure:

$$C(u_T) = \{p \in \Delta(\Omega) : p(T) = 1\}.$$

Proposition 1 (Shapley, 1953) *The set of all unanimity games $\{u_T : T \in \mathcal{F}\}$ is a linear base for $\mathbb{R}^{\mathcal{F}}$. So any $v \in \mathbb{R}^{\mathcal{F}}$ has a unique expression of the form*

$$v = \sum_{T \in \mathcal{F}} \beta_T u_T. \tag{1}$$

Note that, for $E \in \mathcal{F}$, $v(E) = \sum_{T \subseteq E} \beta_T$ and $v'(E) = \sum_{T \cap E \neq \emptyset} \beta_T = \sum_{T \in \mathcal{F}_E} \beta_T$. A game $v \in \mathbb{R}^{\mathcal{F}}$ is said to be *totally monotone* if, in the unique representation (1), β_T is non-negative for all $T \in \mathcal{F}$. A totally monotone capacity is also called a belief function (Dempster, 1967, 1968; Shafer, 1976).

The core operation is additive (in the Minkowski sum) on the set of convex games.[1]

Proposition 2 *If v and w are convex, $C(v) + C(w) = C(v + w)$.*

Proposition 2 can be extended to non-convex games in terms of the Minkowski difference (Danilov and Koshevoy, 2000). The following result is due to Strassen (1964), which is often cited in the literature of belief functions. Technically, it is a simple corollary of the additivity result above, but this turns out to be a very powerful tool.

[1]It is difficult to tell the reference to which this result should be attributed. The result is apparently known in the operations research literature for some time. It is also known in the cooperative game literature.

Corollary 3 (Strassen, 1964) *Let* $v = \sum_T \beta_T u_T$ *be totally monotone. Then,* $p \in C(v)$ *if and only if there exists* $q_T \in \Delta(\Omega)$ *with* $q_T(T) = 1$ *for each* $T \in \mathcal{F}$ *such that* $p = \sum_T \beta_T q_T$. *Equivalently,*

$$C(v) = \sum_{T \in \mathcal{F}} \beta_T C(u_T).$$

2 Dempster-Shafer and maximum likelihood updating

Let v be a belief function and write $v = \sum_T \beta_T u_T$ where $\beta_T \geq 0$. Since $v(\Omega) = \sum_{T \in \mathcal{F}} \beta_T = 1$, the collection $\{\beta_T : T \in \mathcal{F}\}$ defines a probability distribution over $\mathcal{F}$. Then, Corollary 3 suggests the following interpretation of v.

Imagine a decision maker who believes that $\omega \in \Omega$ is chosen by the following two-stage process:

1. an event $T \in \mathcal{F}$ is chosen according to $\{\beta_T : T \in \mathcal{F}\}$,
2. a state $\omega \in T$ is chosen according to some $q_T \in \Delta(\Omega)$ with $q_T(T) = 1$, i.e., $q_T \in C(u_T)$.

Assume that the decision maker knows $\{\beta_T : T \in \mathcal{F}\}$ (and thus v), but does not know q_T for any $T \in \mathcal{F}$. Then, any probability distribution of the form

$$\sum_T \beta_T q_T \text{ where } q_T \in C(u_T)$$

is consistent with the two-stage process. Corollary 3 says that $C(v)$ is exactly the set of all the consistent probability functions.

Now suppose that the decision maker learns from an outside source that an event $E \in \mathcal{F}$ has occurred. There can be many plausible ways to modify the two-stage process in response. The following is certainly one of them.

1. For the first stage, the decision maker should rule out any event which contradicts E: that is, discard any event $T \in \mathcal{F}$ with $T \cap E = \emptyset$, i.e., $T \notin \mathcal{F}_E$.
2. For the second stage, once $T \in \mathcal{F}_E$ has been chosen, the decision maker should rule out any state which contradicts E: that is, discard any sate $\omega \notin T \cap E$.

Dempster and Shafer (Dempster, 1967, 1968; Shafer, 1976) advocated a similar line of reasoning, and showed that the following updating rule captures this idea. The conditional (or updated) capacity given E with $v'(E) > 0$ is defined as follows:

$$v_E^{\mathrm{DS}}(A) := 1 - \frac{v'(A^c \cap E)}{v'(E)} = \frac{v(A \cup E^c) - v(E^c)}{v'(E)} \text{ for all } A \subseteq \Omega, \tag{2}$$

which is often referred to as the Dempster-Shafer (DS) updating rule.

The formula (2) can be reinterpreted utilizing Corollary 3 and the two-stage process interpretation. Fix an event $E \in \mathcal{F}$. The conjugate v' of v gives $v'(E) = \sum_{T\in\mathcal{F}_E} \beta_T$, i.e., $v'(E)$ is the probability that an element in $\mathcal{F}_E$ is chosen in the first stage. Assume throughout that $v'(E) > 0$.

In the modified first stage when $T \in \mathcal{F}_E$ is chosen, since $v'(E) = \sum_{T\in\mathcal{F}_E} \beta_T$, it is natural to assign a probability $\beta_{T|E} := \beta_T/v'(E)$ to each $T \in \mathcal{F}_E$. In the modified second stage when $\omega \in T$ is chosen, the possibility of $\omega \notin T \cap E$ is ruled out. Thus, the decision maker should consider probability functions of the form $\sum_{T\in\mathcal{F}_E} \beta_{T|E} q_T$ where $q_T \in \Delta(\Omega)$ satisfies $q_T(T \cap E) = 1$. These probability functions are exactly those in the elements of $\sum_{T\in\mathcal{F}_E} \beta_{T|E} C(u_{T\cap E})$, which is equal to $C(\sum_{T\in\mathcal{F}_E} \beta_{T|E} u_{T\cap E})$ by Corollary 3. This observation naturally leads us to the following alternative characterization of the DS updating rule:

Theorem 4 *Let $v = \sum_T \beta_T u_T$ be a belief function where $\beta_T \geq 0$ for all $T \in \mathcal{F}$. Then*

$$v_E^{\mathrm{DS}} = \sum_{T\in\mathcal{F}_E} \beta_{T|E} u_{T\cap E}. \tag{3}$$

Proof. For any $A \in \mathcal{F}$,

$$\begin{aligned}
\sum_{T\in\mathcal{F}_E} \beta_{T|E} u_{T\cap E}(A) &= \frac{1}{v'(E)} \sum_{\{T\in\mathcal{F}_E : T\cap E\subseteq A\}} \beta_T \\
&= \frac{1}{v'(E)} \left(\sum_{T\in\mathcal{F}_E} \beta_T - \sum_{\{T\in\mathcal{F}_E : A^c\cap(T\cap E)\neq\emptyset\}} \beta_T \right) \\
&= \frac{1}{v'(E)} \left(\sum_{T\in\mathcal{F}_E} \beta_T - \sum_{\{T : T\cap(A^c\cap E)\neq\emptyset\}} \beta_T \right).
\end{aligned}$$

The last equality holds since $T \cap (A^c \cap E) \neq \emptyset$ implies $T \cap E \neq \emptyset$ and thus $T \in \mathcal{F}_E$. Since $\sum_{T\in\mathcal{F}_E} \beta_T = v'(E)$ and $\sum_{\{T:T\cap(A^c\cap E)\neq\emptyset\}} \beta_T = \sum_{T\in\mathcal{F}_{A^c\cap E}} \beta_T = v'(A^c \cap E)$, the last expression is exactly the same as (2). ∎

Example 1 (The DS updating for unanimity games) Let $v = u_T$. Note that $v'(E) > 0$ if and only if $T \cap E \neq \emptyset$. Since $\beta_{S|E} = 1$ if $S = T$, otherwise $\beta_{S|E} = 0$, we have $v_E^{\mathrm{DS}} = u_{T\cap E}$.

An advantage of thinking the DS updating rule in terms of (3) over (2) is that it is straightforward to show that the DS updating rule is equivalent to the maximum likelihood updating rule over the set of probability functions $C(v)$. For this purpose, fix a belief function $v = \sum_T \beta_T u_T$ and an event $E \in \mathcal{F}$. Define

$$C_E(v) := \left\{ p \in \Delta(\Omega) : p = \sum_{T \in \mathcal{F}} \beta_T q_T,\ q_T \in C(u_{T\cap E}) \text{ if } T \in \mathcal{F}_E,\ q_T \in C(u_T) \text{ otherwise} \right\}.$$

Note that $C_E(v) \subseteq C(v)$ by Corollary 3.

Lemma 5 *$p \in C_E(v)$ if and only if $p \in \arg\max\{p(E) : p \in C(v)\}$. That is, $C_E(v)$ is the set of probability functions in $C(v)$ which maximize the likelihood of E.*

Proof. For any $p = \sum_{T\in\mathcal{F}} \beta_T q_T \in C(v)$ where $q_T \in C(u_T)$,

$$p(E) = \sum_{T\in\mathcal{F}} \beta_T q_T(E) = \sum_{T\in\mathcal{F}_E} \beta_T q_T(E) \leq \sum_{T\in\mathcal{F}_E} \beta_T.$$

The equality holds if and only if, for every $T \in \mathcal{F}_E$ with $\beta_T > 0$, $q_T(E) = 1$, i.e., $q_T \in C(u_{T\cap E})$. Since $\sum_{T\in\mathcal{F}} \beta_T q_T = \sum_{T\in\mathcal{F}} \beta_T q'_T$ as far as $q_T = q'_T$ for $T \in \mathcal{F}$ with $\beta_T > 0$, the lemma is proved. ∎

Now we shall relate $C_E(v)$ to the DS updating rule: the DS updating rule is equivalent to the maximum likelihood updating rule over $C(v)$. This result is reported in Gibloa and Schmeidler (1993) and Denneberg (1994), with different proofs.

Proposition 6 *$q \in C(v_E^{\mathrm{DS}})$ if and only if $q = p(\cdot|E)$ for some $p \in C_E(v)$.*

Proof. Pick any $p = \sum_{T\in\mathcal{F}} \beta_T q_T \in C_E(v)$. Then, for each $A \in \mathcal{F}$, we have:

$$p(A\cap E) = \sum_{T\in\mathcal{F}} \beta_T q_T(A\cap E) = \sum_{T\in\mathcal{F}_E} \beta_T q_T(A\cap E) = \sum_{T\in\mathcal{F}_E} \beta_T q_T(A).$$

Indeed, the first equation holds by definition, and the second holds since $q_T(A\cap E) \leq q_T(E) = 0$ for every $T \notin \mathcal{F}_E$. For the last, note that $q_T(A) \geq q_T(A\cap E) \geq q_T(A\cap T\cap E)$. But if $T \in \mathcal{F}_E$ then $q_T \in C(u_{T\cap E})$ and $q_T(A) = q_T(A\cap T\cap E)$, which implies $q_T(A) = q_T(A\cap E)$.

Notice in particular that $p(E) = \sum_{T\in\mathcal{F}_E} \beta_T q_T(E) = \sum_{T\in\mathcal{F}_E} \beta_T = v'(E)$. Thus, for any A,

$$p(A|E) = p(A\cap E)/p(E) = \frac{1}{v'(E)} \sum_{T\in\mathcal{F}_E} \beta_T q_T(A) = \sum_{T\in\mathcal{F}_E} \beta_{T|E} q_T(A).$$

Thus, $p(\cdot|E) \in \sum_{T\in\mathcal{F}_E} \beta_{T|E} C(u_{T\cap E}) = C(\sum_{T\in\mathcal{F}_E} \beta_{T|E} u_{T\cap E}) = C(v_E^{\mathrm{DS}})$ from (3).

On the other hand, start with any $q \in C(v_E^{\mathrm{DS}})$, thus we can write $q = \frac{1}{v'(E)} \sum_{T\in\mathcal{F}_E} \beta_T q_T$ with $q_T \in C(u_{T\cap E})$ by (3). Fix $q_T \in C(u_T)$ arbitrarily for $T \notin \mathcal{F}_E$, and define $\hat{p} := \sum_{T\in\mathcal{F}} \beta_T q_T$. Then $\hat{p} \in C_E(v)$ by construction, and $q = \hat{p}(\cdot|E)$. ∎

Example 2 (The DS updating for unanimity games, ctd.) If $v = u_T$ and $T \cap E \neq \emptyset$, $p \in C_E(v)$ if and only if $p(T \cap E) = 1$, which is equivalent to $p \in C(u_{T \cap E})$.

References

Danilov, V. I., and Koshevoy, G. A. (2000). "Cores of Cooperative Games, Superdifferentials of Functions, and the Minkowski Difference of Sets," *Journal of Mathematical Analysis and Applications* 247, 1–14.

Dempster, A. P. (1967). "Upper and Lower Probabilities Induced by a Multi-valued Mapping," *Annals of Mathematical Statistics* 38, 325–339.

Dempster, A. P. (1968). "A Generalization of Bayesian Inference," *Journal of Royal Statistical Society, Series B* 38, 205–247.

Denneberg, D. (1994). "Conditioning (Updating) Non-additive Measures," *Annals of Operations Research* 52, 21–42.

Gilboa, I., and Schmeidler, D. (1993). "Updating Ambiguous Beliefs," *Journal of Economic Theory* 59, 33–49.

Shafer, G. (1976). *A Mathematical Theory of Evidence.* Princeton, NJ: Princeton Univ. Press.

Shapley, L. S. (1953). "A Value for n-Person Games," in *Contributions to the Theory of Games* II (H. Kuhn and A. Tucker, Eds.), pp. 307–317, Princeton, NJ: Princeton Univ. Press.

Strassen, V. (1964). "Messfehler und Information," *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete* 2, 273–305.