UDC 519.816::519.812.3

Minimaximax approach for finding optimal decisions' subset regarding changes of the loss function

V. V. Romanuke
Khmelnitskiy National University, Ukraine

A generalization of the decision (loss or utility) function is suggested. An ordinary decision function is defined on a Cartesian product of a decisions' set and a set of states, but the generalized decision function has the third variable called a metastate. Metastates are generated due to uncertain evaluation of ordinary situations, or influence of the time course. For minimizing losses under poor or unreliable statistics, the rule of minimaximax is fully described. For correctly transferring from minimaximax to Bayesian criterions, the rules of minimizing expected losses for the generalized loss function are formalized. All the suggested criterions are re-formalized for the case of the utility function.

Key words: *decision, minimax, metastate, loss function, minimaximax.*

Пропонується узагальнення функції рішень (втрат або корисності). Звичайна функція рішень визначається на декартовому добутку множини рішень та множини станів, тоді як узагальнена функція рішень має третю змінну, котра називається метастаном. Метастани породжуються внаслідок невизначеного оцінювання звичайних ситуацій або впливу плину часу. Для мінімізації втрат за слабкої або ненадійної статистики ґрунтовно описується правило мінімаксимаксу. Для коректного переходу від мінімаксимаксу до байєсових критеріїв правила мінімізації очікуваних втрат формалізуються для узагальненої функції втрат. Усі запропоновані критерії переформалізовуються для випадку функції корисності.

Ключові слова: *рішення, мінімакс, метастан, функція втрат, мінімаксимакс.*

Предлагается обобщение функции решений (потерь или полезности). Обычная функция решений определяется на декартовом произведении множества решений и множества состояний, тогда как обобщённая функция решений имеет третью переменную, называемую метасостоянием. Метасостояния порождаются вследствие неопределённого оценивания обычных ситуаций или влияния течения времени. Для минимизации потерь при слабой или ненадёжной статистике основательно описывается правило минимаксимакса. Для корректного перехода от минимаксимакса к байесовым критериям правила минимизации ожидаемых потерь формализируются для обобщённой функции потерь. Все предложенные критерии переформализируются для случая функции полезности.

Ключевые слова: *решение, минимакс, метасостояние, функция потерь, минимаксимакс.*

Inconstancy of the loss function

The loss function is an important mathematical object relating to a wide variety of technical, economical, and social processes [1, 2]. An antipode to the loss function is the utility function which is of much rarer application [1, 3]. However, this function may severely change, especially if a process is studied for a longer period and requires more reliable optimality of decisions. Another issue is poor statistics, not allowing to evaluate each pair of a decision and a state with a point value [4, 5]. A simple example is a risk matrix whose elements take on a few possible values due to expert judgments/estimations. Therefore, finding an optimal decisions' subset should regard changes of the loss function.

Approaches to minimizing losses under poor or unreliable statistics

A great number of approaches and heuristics exists for finding an optimal decisions' subset by a known loss/utility function [1, 6]. When statistics is poor or unreliable, losses are minimaxed for ensuring minimal damage under worst possibly conditions [7]. Then the criterion of Wald or the criterion of Savage is applied [1, 6, 7]. In the article [8], for ensuring industrial and manufacturing labor safety, a meta-minimax approach has been represented regarding the change of the risk matrix. The change has been explained due to the impossibility of the point evaluation of the risk matrix and therefore this matrix has been represented as a finite set of matrices, implying the risk matrix change through that set. Each version of the risk matrix has been tied to a state which is called a metastate. Thus the finite change of the risk matrix has been substituted with the three-dimensional risk matrix. The article [8] suggests the finite minimaximax along with another three cases of minimizing the risk. These cases admit availability of statistics:

1) probabilistic measures relating to ordinary states are available;

2) a probabilistic measure over metastates is available;

3) both are available.

Factually, the three-dimensional risk matrix is equivalent to a finite series of decision making problems (or a multiple state decision making problem). In the article [9], an algorithm of reducing a finite series of decision making problems to a single problem has been suggested. But, without statistics, it works only if there is a nonempty intersection of the optimal decisions' subsets for the metastates.

Goal of the article and the tasks to be accomplished

In view of the fact that neither the article [8] nor the article [9] contain full description of minimaximax, the goal is to fully describe the minimaximax criterion for the loss function with metastates (changes). For transferring from minimaximax to criterions of minimizing expected losses, application of probabilistic measures will be described. For reaching the goal, the following tasks are to be accomplished:

1. To formalize changes of the loss function.

2. To formalize the rule of minimaximax for finding an optimal decisions' subset regarding the changes. For appropriate formalization, both finite and infinite cases will be considered for the sets of decisions, states, metastates.

3. To formalize the rules of minimizing expected losses for the loss function with metastates. Application of probabilistic measures will be described for the three cases:

3.1. Probabilistic measures over ordinary states are available, and a probabilistic measure over metastates is unavailable.

3.2. Probabilistic measures over ordinary states are unavailable, and a probabilistic measure over metastates is available.

3.3. Probabilistic measures of both types are available.

4. To re-formalize all the suggested criterions for the case of the utility function.

Formalization of the loss function's changes

Let X be a set of decisions, and S be a set of states. If X is finite then $X = \{x_i\}_{i=1}^{N}$ by $N = |X|$ and $N \in \mathbb{N} \setminus \{1\}$. Similarly, $S = \{s_j\}_{j=1}^{Q}$ by $Q = |S|$ and

$Q \in \mathbb{N} \setminus \{1\}$ for a finite set S. Denote the set of metastates by M. Each element of this set implies existence of its own loss function. When a metastate shifts to other one, the loss function changes. If M is finite then $M = \{m_k\}_{k=1}^{K}$ by $K = |M|$ and $K \in \mathbb{N} \setminus \{1\}$.

Let a real value $r(x, s, m)$ be a loss (risk) in the situation

$$\{x, s, m\} \text{ by } x \in X, \ s \in S, \ m \in M. \quad (1)$$

Therefore, $r(x, s, m)$ is a (generalized) loss function defined on the set $X \times S \times M$.

Definition 1. The loss function $r(x, s, m)$ is called finite if the set $X \times S \times M$ is finite. The finite loss function $r(x, s, m)$ is called the generalized loss matrix.

Factually, the generalized loss matrix is a stack of K ordinary $N \times Q$ matrices. Let $r_{ijk} = r(x_i, s_j, m_k)$ in the situation $\{x_i, s_j, m_k\}$. Then $\mathbf{R}_k = (r_{ijk})_{N \times Q}$ is an ordinary loss matrix at the k-th metastate, and the generalized loss matrix is a set $\{\mathbf{R}_k\}_{k=1}^{K}$. Formally, finite changes of a decision matrix can be substituted with a loss $N \times Q \times K$ matrix.

Definition 2. The loss function $r(x, s, m)$ is called countable if the set $X \times S \times M$ is countable.

Definition 3. The loss function $r(x, s, m)$ is called infinite if the set $X \times S \times M$ is infinite.

It is easy to see that if just one of the sets X, S, M is infinite then the loss function is infinite. A countably infinite loss function cannot be represented as a stack of matrices. Properties of infinite countability among sets X, S, M do not necessarily coincide.

Definition 4. The loss function $r(x, s, m)$ is called continuous if the set $X \times S \times M$ is continuous.

To be continuous, each of the sets X, S, M must be continuous. The loss function becomes discontinuous if just one of the sets X, S, M has a discontinuity.

Rule of minimaximax

When any probabilistic measures are unavailable or uncertain, the classical minimax approach [1, 6, 7] guarantees the optimal loss, although the most pessimistic. The pessimism concerns all the noncontrollable states, i. e. metastates as well. Then, the optimal decisions' subset $X^* \subset X$ for the loss function $r(x, s, m)$ is found by the minimaximax rule as

$$X^* = \arg\min_{x \in X} \left\{ \max_{s \in S} \max_{m \in M} r(x, s, m) \right\} \subset X. \quad (2)$$

Obviously, if X is infinite, X^* can have infinitely many optimal decisions. Whether each set of decisions, states, metastates is finite or infinite, statement (2) for minimaximax remains correct. For the generalized loss matrix, minimaximax by (2) is re-written simpler [8]:

$$X^* = \arg \min_{x_i,\ i=\overline{1,\ N}} \left\{ \max_{j=\overline{1,\ Q}} \max_{k=\overline{1,\ K}} r_{ijk} \right\} \subset X \ . \tag{3}$$

Minimaximax by (2) or (3) is an effective rule when conditions, under which one has to make a decision, occur rarely or just a few times [1, 4, 5, 8, 9]. Although losses are maximized twice, severity of the minimaximax is not bigger than that of the classical minimax approach. This is explained with that, for both rules, pessimism is directed against noncontrollability.

Minimizing expected losses for the generalized loss function

Suppose that, for a finite loss function, probabilistic measures relating to ordinary states are available. Availability here relates to reliability of point-valued probabilistic estimations. Let $\mathbf{P}_k = \left(p_{ijk} \right)_{N \times Q}$ be the stochastic matrix whose value p_{ijk} is the probability of the j-th state at the k-th metastate when the i-th decision is selected. It is obvious that

$$\sum_{j=1}^{Q} p_{ijk} = 1 \quad \forall\ i = \overline{1,\ N} \quad \text{and} \quad \forall\ k = \overline{1,\ K} \ . \tag{4}$$

Then the finite subset [8]

$$X^* = \arg \min_{x_i,\ i=\overline{1,\ N}} \left\{ \max_{k=\overline{1,\ K}} \sum_{j=1}^{Q} p_{ijk} r_{ijk} \right\} \subset X \tag{5}$$

contains decisions which minimize the maximally expected losses with respect to ordinary states.

Generally, let $p(x,\ s,\ m)$ be a probabilistic measure over ordinary states for each decision $x \in X$ and each metastate $m \in M$. The non-negative function $p(x,\ s,\ m)$ is defined on the set S with a Lebesgue measure $\mu_S(s)$, so

$$\int_S p(x,\ s,\ m)\, d\mu_S(s) = 1 \quad \forall\ x \in X \quad \text{and} \quad \forall\ m \in M \ . \tag{6}$$

Then the maximally expected losses with respect to ordinary states are minimized:

$$X^* = \arg \min_{x \in X} \left\{ \max_{m \in M} \int_S p(x,\ s,\ m)\, r(x,\ s,\ m)\, d\mu_S(s) \right\} \subset X \ . \tag{7}$$

The integration over the finite set S in (7) is substituted with the summation:

$$X^* = \arg \min_{x \in X} \left\{ \max_{m \in M} \sum_{j=1}^{Q} p(x,\ s_j,\ m)\, r(x,\ s_j,\ m) \right\} \subset X \ , \tag{8}$$

where $p(x, s_j, m)$ is the probability of the j-th state at metastate $m \in M$ when the decision $x \in X$ is selected, and

$$\sum_{j=1}^{Q} p(x, s_j, m) = 1 \quad \forall\, x \in X \quad \text{and} \quad \forall\, m \in M\,. \tag{9}$$

The optimal decisions' subset by (7), or its equivalents (5) and (8), is effective when conditions under which one has to make a decision occur frequently [1, 8, 9]. Application of the formula (8) is more likely, because finite point estimation is far more reliable than infinite one [4, 8].

An easier case is when probabilistic measures over ordinary states are unavailable, but a probabilistic measure over metastates is available. Suppose that, for a finite loss function, probabilities $\{w_k\}_{k=1}^{K}$ are known, where w_k is the probability of the k-th metastate and

$$\sum_{k=1}^{K} w_k = 1\,. \tag{10}$$

Then the finite subset [8]

$$X^* = \arg \min_{x_i,\, i=\overline{1,\, N}} \left\{ \max_{j=\overline{1,\, Q}} \sum_{k=1}^{K} w_k r_{ijk} \right\} \subset X \tag{11}$$

contains decisions which minimize the maximally expected losses with respect to metastates. Minimax by (11) reminds the classic minimax applied to a decision matrix (of losses, regrets, or risks).

Generally, let $w(m)$ be a probabilistic measure over metastates. The nonnegative function $w(m)$ is defined on the set M with a Lebesgue measure $\mu_M(m)$, so

$$\int_M w(m)\, d\mu_M(m) = 1\,. \tag{12}$$

Then the maximally expected losses with respect to metastates are minimized:

$$X^* = \arg\min_{x \in X} \left\{ \max_{s \in S} \int_M w(m)\, r(x, s, m)\, d\mu_M(m) \right\} \subset X\,. \tag{13}$$

The integration over the finite set M in (13) is substituted with the summation:

$$X^* = \arg\min_{x \in X} \left\{ \max_{s \in S} \sum_{k=1}^{K} w_k r(x, s, m_k) \right\} \subset X\,. \tag{14}$$

Abstracting from the values under maxima, (13) and (14) both coincide with the classic minimax.

The optimal decisions' subset by (13), or its equivalents (11) and (14), is effective when the loss function changes (over metastates) frequently [1, 2, 6, 8, 9]. Although evaluation of probabilities $\{w_k\}_{k=1}^{K}$ or the probabilistic measure $w(m)$ is a non-trivial separate problem, availability of them is much more likely than that of probabilistic measures in matrices $\{\mathbf{P}_k\}_{k=1}^{K}$ or functions $p(x, s, m)$ for each decision and each metastate. That is why minimax by (11), (14), or (13) is much more practical than minimax by (5), (8), (7), respectively (the formulas are listed in order of increasing complexity).

The case when all the probabilistic measures are available is more theoretical rather than practical one. In the simplest case, which is the finite one,

$$X^* = \arg \min_{x_i,\, i=\overline{1, N}} \left\{ \sum_{k=1}^{K} w_k \sum_{j=1}^{Q} p_{ijk} r_{ijk} \right\} \subset X . \tag{15}$$

Another three cases of the finite/infinite probabilistic measures produce the following formulas:

$$X^* = \arg\min_{x \in X} \left\{ \int_M w(m) \sum_{j=1}^{Q} p(x, s_j, m) r(x, s_j, m) d\mu_M(m) \right\} \subset X , \tag{16}$$

$$X^* = \arg\min_{x \in X} \left\{ \sum_{k=1}^{K} w_k \int_S p(x, s, m_k) r(x, s, m_k) d\mu_S(s) \right\} \subset X , \tag{17}$$

$$X^* = \arg\min_{x \in X} \left\{ \int_M w(m) \int_S p(x, s, m) r(x, s, m) d\mu_S(s) d\mu_M(m) \right\} \subset X . \tag{18}$$

For infinite measures, the optimal decisions' subset by (18) is very impracticable. The subset by either (16) or (17) keeps theory also, unless the measures are defined on countable sets. The case (15) is plausible, though.

The utility function case

Decision making practices utility functions much less than loss or risk functions. Nevertheless, a lot of branches of economic operate on profitability, benefit, gain, revenue, etc.

For the case of the utility function $u(x, s, m)$, the rule of minimaximax becomes the rule of maximinimin. In this way, formulas (2) and (3) are:

$$X^* = \arg\max_{x \in X} \left\{ \min_{s \in S} \min_{m \in M} u(x, s, m) \right\} \subset X , \tag{19}$$

$$X^* = \arg \max_{x_i,\, i=\overline{1, N}} \left\{ \min_{j=\overline{1, Q}} \min_{k=\overline{1, K}} u_{ijk} \right\} \subset X , \tag{20}$$

where $\mathbf{U}_k = \left(u_{ijk}\right)_{N\times Q}$ is an ordinary utility matrix at the k-th metastate.

Maximization of the expected utility is fulfilled by formulas which are symmetrical to formulas (5), (7), (8), (11), (13) — (18), respectively:

$$X^* = \arg \max_{x_i,\ i=\overline{1,\ N}} \left\{ \min_{k=\overline{1,\ K}} \sum_{j=1}^{Q} p_{ijk} u_{ijk} \right\} \subset X \ , \tag{21}$$

$$X^* = \arg\max_{x\in X} \left\{ \min_{m\in M} \int_S p(x,\ s,\ m) u(x,\ s,\ m) d\mu_S(s) \right\} \subset X \ , \tag{22}$$

$$X^* = \arg\max_{x\in X} \left\{ \min_{m\in M} \sum_{j=1}^{Q} p(x,\ s_j,\ m) u(x,\ s_j,\ m) \right\} \subset X \ , \tag{23}$$

$$X^* = \arg \max_{x_i,\ i=\overline{1,\ N}} \left\{ \min_{j=\overline{1,\ Q}} \sum_{k=1}^{K} w_k u_{ijk} \right\} \subset X \ , \tag{24}$$

$$X^* = \arg\max_{x\in X} \left\{ \min_{s\in S} \int_M w(m) u(x,\ s,\ m) d\mu_M(m) \right\} \subset X \ , \tag{25}$$

$$X^* = \arg\max_{x\in X} \left\{ \min_{s\in S} \sum_{k=1}^{K} w_k u(x,\ s,\ m_k) \right\} \subset X \ , \tag{26}$$

$$X^* = \arg \max_{x_i,\ i=\overline{1,\ N}} \left\{ \sum_{k=1}^{K} w_k \sum_{j=1}^{Q} p_{ijk} u_{ijk} \right\} \subset X \ , \tag{27}$$

$$X^* = \arg\max_{x\in X} \left\{ \int_M w(m) \sum_{j=1}^{Q} p(x,\ s_j,\ m) u(x,\ s_j,\ m) d\mu_M(m) \right\} \subset X \ , \tag{28}$$

$$X^* = \arg\max_{x\in X} \left\{ \sum_{k=1}^{K} w_k \int_S p(x,\ s,\ m_k) u(x,\ s,\ m_k) d\mu_S(s) \right\} \subset X \ , \tag{29}$$

$$X^* = \arg\max_{x\in X} \left\{ \int_M w(m) \int_S p(x,\ s,\ m) u(x,\ s,\ m) d\mu_S(s) d\mu_M(m) \right\} \subset X \ . \tag{30}$$

Conditions of effectiveness for subsets (19) — (30) are similar to those for subsets (2), (3), (5), (7), (8), (11), (13) — (18).

Arrangement of the ordinary loss functions into the generalized loss function

Metastates are generated due to that situations $\{x, s\}$ are usually evaluated with intervals. When the intervals are sampled, a finite set of metastates is produced. Another reason that causes metastates' generation is the time course. Thus, in practice, point estimation or evaluation of the loss function is impossible. Then, the generalized loss function is represented as the set M of the ordinary loss functions. It implies the loss function changes through this set [2, 9]. Arranging its elements is a pretty hard problem, which is explicitly solved in [8] for the generalized loss matrix only.

If $\mathbf{R}_1$ is initially given, then the following $K-2$ matrices satisfy the condition [8]:

$$\mathbf{R}_k \in \arg \min_{\mathbf{R}_h,\, h=\overline{k,\, K}} \rho_{N\times Q}\left(\mathbf{R}_{k-1},\, \mathbf{R}_h\right) \text{ for } k=\overline{2,\, K-1} \tag{31}$$

by the distance $\rho_{N\times Q}\left(\mathbf{A},\, \mathbf{B}\right)$ in the space of real-valued $N\times Q$ matrices $\mathbf{A}$ and $\mathbf{B}$. Matrix $\mathbf{R}_K$ remains itself after matrices $\left\{\mathbf{R}_k\right\}_{k=1}^{K-1}$ are already arranged (by their indices). The condition (31) ensures "resemblance" between neighboring indexed matrices $\mathbf{R}_k$ and $\mathbf{R}_{k+1}$ by $k=\overline{1,\, K-1}$. Such resemblance (by indexed metastates) can be made by arranging matrices $\left\{\mathbf{P}_k\right\}_{k=1}^{K}$ in the same way, when $\mathbf{P}_1$ is initially given [8]. Then, however, "resemblance" between neighboring indexed matrices $\mathbf{P}_k$ and $\mathbf{P}_{k+1}$ is not necessarily accompanied with "resemblance" between neighboring indexed matrices $\mathbf{R}_k$ and $\mathbf{R}_{k+1}$, $k=\overline{1,\, K-1}$.

Conclusion

The suggested and fully described minimaximax approach in (2), (3), or maximinimin in (19), (20), allows to take into account changes of the decision (loss or utility) functions. It is necessary because even the best-assurance minimax/maximin criterion turns out to be inconsistent for uncertain evaluations of a situations $\{x, s\}$. Meta-situation (1) expresses that uncertainty. Besides, it regards the course of time. If a situations $\{x, s\}$ are evaluated by different experts without consensus, this also results in meta-situations.

It is important that the set of metastates (changes) is not just finite, but can be countable, infinite, or continuous/discontinuous. Furthermore, the optimality by (5), (7), (8), (11), (13) — (18), and (21) — (30) is presented for correctly transferring from minimaximax to criterions of improving expectations (Bayesian criterions). But arranging the ordinary decision functions for the set of metastates is an open question concerning both minimaximax and criterions with expectations. Selection of an initial loss function and an initial bunch of probabilistic measures on the set $X\times S$, similar to $\mathbf{R}_1$ and $\mathbf{P}_1$, is discussible as well.

REFERENCES

1. Трухаев Р. И. Модели принятия решений в условиях неопределённости / Трухаев Р. И. — М. : Наука, 1981. — 258 с.
2. Lark R. M. The implicit loss function for errors in soil information / R. M. Lark, K. V. Knights // Geoderma. — 2015. — Vol. 251 — 252. — P. 24 — 32.
3. Biederman D. K. A strictly-concave, non-spliced, Giffen-compatible utility function / D. K. Biederman // Economics Letters. — 2015. — Vol. 131. — P. 24 — 28.
4. Li Y. P. A robust interval-based minimax-regret analysis approach for the identification of optimal water-resources-allocation strategies under uncertainty / Y. P. Li, G. H. Huang, S. L. Nie // Resources, Conservation and Recycling. — 2009. — Vol. 54, Iss. 2. — P. 86 — 96.
5. Dong C. An interval-parameter minimax regret programming approach for power management systems planning under uncertainty / C. Dong, G. H. Huang, Y. P. Cai, Y. Xu // Applied Energy. — 2011. — Vol. 88, Iss. 8. — P. 2835 — 2845.
6. Information, Inference and Decision / Ed. by G. Menges. — Dordrecht : D. Reidel Publishing Company, 1974. — 201 p.
7. Moon J. Minimax estimation with intermittent observations / J. Moon, T. Başar // Automatica. — 2015. — Vol. 62. — P. 122 — 133.
8. Romanuke V. V. Meta-minimax approach for optimal alternatives subset regarding the change of the risk matrix in ensuring industrial and manufacturing labor safety / V. V. Romanuke // Herald of Khmelnytskyi national university. Technical sciences. — 2015. — № 6. — P. 97 — 99.

Romanuke V. V. Multiple state problem reduction and decision making criteria hybridization / V. V. Romanuke // Research Bulletin of NTUU “Kyiv Polytechnic Institute”. — 2016. — № 2. — С. 51 — 59.

Надійшла у першій редакції 15.05.2017, в останній – 22.05.2017