Aversion for Imprecise Probabilities in Linear Programming

By

Somdeb Lahiri

(email: somdeb.lahiri@gmail.com)

ORCID: 0000-0002-5247-3497

(Formerly with) PD Energy University, Gandhinagar (EU-G), India.

November 26, 2026.

Abstract

In this paper we recall the well-known results that an alternative is weakly efficient if and only if it is a constrained expected utility maximizer and efficient if and only if it is an expected utility maximizer with respect to a probability distribution having full support. We provide simple direct proofs for the two results using Farkas' lemma. Our main result is the formulation of the problem of finding a max-min expected pay-off maximizer under ambiguity (with imprecise probabilities) as a linear programming problem.

AMS Classifications: 90C05, 90C29, 91B06

JEL Classifications: C61, D81

Keywords: multi-objective linear programming, ambiguity, imprecise probabilities, max-min, linear programming

1. Introduction: In decision-making under uncertainty, *risk* is the situation where the decision-maker has precise probabilistic information about the realization of possible uncertain states of nature (events) exactly one of which is going to be realized at a future date. The absence of such precise probabilistic information is what is generally referred to as (informational) *ambiguity*. The investigation of *decision-making under ambiguity* as a stand-alone scholarly purpose originates in the work of Ellsberg (1961). In a *multi-objective linear programming problem (MOLP)* the decision maker is faced with (the possibility of) multiple linear objective functions where choices are restricted to a feasible set of alternatives defined by linear constraints. A focused and lucid discussion of multi-objective linear programming can be found in Luptacik (2010).

Our purpose here is to view multi-objective linear programming from the perspective of decision-making under ambiguity. We assume that there is a non-empty finite set of states of nature, exactly one of which will be realized at a future date. Corresponding to each possible future state of nature, there is a linear pay-off function

defined on the set of all alternatives. Each alternative is a point in a finite dimensional Euclidean space. If an alternative yields a strictly greater pay-off than another alternative, regardless of the state of nature that is realized, there is no reason why the decision maker would choose the second alternative. An alternative whose state-dependent pay-off vector is *not strictly less* than the state-dependent pay-off vector of another alternative, is said to be *weakly efficient*. A stronger version of efficiency for any alternative requires that its state-dependent pay-off vector is *not less than or equal to and unequal* to the pay-off vector of any other alternative. In this context we prove two well-known results:

(A) An alternative is efficient if and only if there exists a probability distribution with full support over the states of nature such that the alternative maximizes expected pay-off subject to the given linear constraints.

(B) An alternative is weakly efficient if and only if there exists a probability distribution (with *not necessarily full support*) over the states of nature such the alternative maximizes expected pay-off subject to the given linear constraints.

Unlike the indirect proofs of these two "scalarization" results available in Luc (2016), we provide direct proofs of the two results using Farka's lemma. The implication of these two results is that an alternative is consistent with "Bayesian rationality" if and only if it is efficient, either weakly or strongly. However, it says nothing about which efficient alternative would or should be chosen if there is ambiguity.

We present a very general model of ambiguity that merely requires the probability of the occurrence of a state of nature to belong to a (possibly) state-dependent interval of real numbers. This is the case with what is commonly known as "imprecise probabilities". Having done so, we assume that the decision-maker is pessimistic in the sense that it evaluates each alternative by considering the worst expected pay-off of the alternative and then chooses an alternative with the highest such evaluation. This is known as the max-min rule, a brief, lucid and precise discussion of which is available in pages 171-172 of Gilboa (2011). A rigorous axiomatic characterization of the max-min rule is available in Gilboa and Schmeidler (1989). The essence of the max-min rule is "*ambiguity aversion*".

We can now transform the problem faced by the decision-maker into a two-person zero sum (matrix) game against nature, where the set of pure strategies for the decision maker is the non-empty, finite set of extreme points of the set of all feasible alternatives available to it and the set of pure strategies for nature is the set of extreme

points of the set of imprecise probability vectors available to the latter. The seminal work of George Dantzig (Dantzig (1951)) proves that a randomization is a max-min solution of a matrix game if and only if it solves a related linear programming problem. Thus, an alternative is a max-min solution for the problem faced by the decision-maker if and only if there exists a randomization that solves a certain linear programming problem, and the alternative itself is the weighted sum of the extreme points of the feasible set where the weights are as in the solution of the linear programming problem. On the way, we provide an example (in note 4) an implication of which is the well known fact, that the theory of linear programming is a *strict generalization* of matrix game theory.

A linear programming formulation of resolving ambiguity using the optimism-pessimism rule (i.e., Hurwicz criterion) is available in Lahiri (2026).

2. The Model: For positive integers m, n, p, let $C \epsilon \mathbb{R}^{p \times n}$, $A \epsilon \mathbb{R}^{m \times n}$ and $b \epsilon \mathbb{R}^{m}$.

For any matrix, a subscript denotes a row of the matrix and a super-script its column. Superscript "T" is used to denote transpose.

Consider the multi-objective linear programming problem: Maximize Cx, subject to $Ax = b$, $x \epsilon \mathbb{R}^{n}_{+}$.

Let $X(A, b) = \{x \epsilon \mathbb{R}^{n}_{+} | Ax = b\}$.

$x^* \epsilon X(A, b)$ is said to be **efficient** if there does not exist $x \epsilon X(A, b)$ satisfying $Cx > Cx^*$ (i.e., $C_h x \geq C_h x^*$ for all $h \epsilon \{1, \ldots, p\}$, $Cx \neq Cx^*$).

$x^* \epsilon X(A, b)$ is said to be **weakly efficient** if there does not exist $x \epsilon X(A, b)$ satisfying $Cx >> Cx^*$ (i.e., $C_h x > C_h x^*$ for all $h \epsilon \{1, \ldots, p\}$).

Note 1: If for some $h \epsilon \{1, \ldots, p\}$, $C_h x = C_h x^*$, then for all $t > 0$ it must be the case that $tC_h x = tC_h x^*$. If for some $h \epsilon \{1, \ldots, p\}$, $C_h x > C_h x^*$, then for all $t > 0$ it must be the case that $tC_h x > tC_h x^*$. ▌

3. Scalarization For Efficient Alternatives:

Theorem 1: x^* is efficient if and only if there exists $\lambda \epsilon \mathbb{R}^{p}_{++}$ such that $\lambda^T C x^* \geq \lambda^T C x$ for all $x \epsilon X(A, b)$.

Proof: Let $I(x^*) = \{i | x_i^* = 0\}$.

x^* is efficient if and only if there <u>does not exist</u> $h \epsilon \mathbb{R}^n$ satisfying $h_j \geq 0$ for all $j \epsilon I(x^*)$, $Ah = 0$ and $Ch > 0$.

Thus, x^* is efficient if and only if, there does not exist $\langle h_j | j \epsilon I(x^*) \rangle$, $\langle h_j^{(1)} | j \notin I(x^*) \rangle$, $\langle h_j^{(2)} | j \notin I(x^*) \rangle$ all non-negative and $g \epsilon \mathbb{R}^{p}_{+}$ such that:

$\sum_{j\in I(x^*)} A^j h_j + \sum_{j\notin I(x^*)} A^j h_j^{(1)} - \sum_{j\notin I(x^*)} A^j h_j^{(2)} = 0,$

$\sum_{j\in I(x^*)} C^j h_j + \sum_{j\notin I(x^*)} C^j h_j^{(1)} - \sum_{j\notin I(x^*)} C^j h_j^{(2)} - I^{(p)}g = 0$

$E^{(p)T}g = 1.$

By Farkas' lemma, x^* is efficient if and only if there exists $\xi\epsilon\mathbb{R}^m$, $\lambda\epsilon\mathbb{R}^p$, $\theta\epsilon\mathbb{R}$ such that $\xi^T A^j + \lambda^T C^j \leq 0$ for all $j\epsilon\{1, \ldots, n\}$, $\xi^T A^j + \lambda^T C^j \leq 0$ for all $j\notin I(x^*)$, $-\lambda^T + \theta E^{(p)T} \leq 0$, $\theta > 0$.

Thus, x^* is efficient if and only if there exists $\xi\epsilon\mathbb{R}^m$, $\lambda\epsilon\mathbb{R}^p$ such that $\xi^T A^j + \lambda^T C^j \leq 0$ for all $j\epsilon I(x^*)$, $\xi^T A^j + \lambda^T C^j = 0$ for all $j\notin I(x^*)$, $\lambda >> 0$.

Thus, if x^* is efficient then there exists $\lambda\epsilon\mathbb{R}^p_{++}$ such that for all $h\epsilon\mathbb{R}^n$ satisfying $h_j \geq 0$ for all $j\epsilon I(x^*)$, $Ah = 0$, it is the case that $\lambda^T Ch \leq 0$.

For $h\epsilon\mathbb{R}^n$, $x^* + th \epsilon X(A,b)$ for all $t > 0$ sufficiently small if and only if [$h_j \geq 0$ for all $j\epsilon I(x^*)$ and $Ah = 0$].

Thus, if x^* is efficient then there exists $\lambda\epsilon\mathbb{R}^p_{++}$ such that $\lambda^T Cx^* \geq \lambda^T Cx$ for all $x\epsilon X(A, b)$.

Now suppose there exists $\lambda\epsilon\mathbb{R}^p_{++}$ such that $\lambda^T Cx^* \geq \lambda^T Cx$ for all $x\epsilon X(A, b)$.

Towards a contradiction suppose there exists $x\epsilon X(A, b)$ such that $Cx > Cx^*$.

Since $\lambda\epsilon\mathbb{R}^p_{++}$, $Cx > Cx^*$ implies $\lambda^T Cx > \lambda^T Cx^*$, contrary to our assumption.

Thus, x^* must be efficient. Q.E.D.

Note 2: The interpretation of theorem 1 in our context is that x^* is efficient <u>if and only if</u> there exists a probability distribution over $\{1, \ldots, p\}$ with respect to which the decision maker maximizes the expected value of the objective function subject to the given linear constraints. ▌

As an immediate consequence of theorem 1 and duality theory of linear programming we have the following corollary.

Corollary 1 of Theorem 1: $x^*\epsilon X(A, b)$ is efficient if and only if there exists $\lambda\epsilon\mathbb{R}^p_{++}$ and $y^*\epsilon\mathbb{R}^m$ such that $y^{*T}A - \lambda^T C \geq 0$ and $(y^{*T}A - \lambda^T C)x^* = 0$.

Theorem 2: x^* is weakly efficient if and only if there exists $\lambda\epsilon\mathbb{R}^p_+\backslash\{0\}$ such that $\lambda^T Cx^* \geq \lambda^T Cx$ for all $x\epsilon X(A, b)$.

Proof: Let $I(x^*) = \{i | x_i^* = 0\}$.

x^* is weakly efficient if and only if there <u>does not exist</u> $h\epsilon\mathbb{R}^n$ satisfying $h_j \geq 0$ for all $j\epsilon I(x^*)$, $Ah = 0$ and $Ch >> 0$.

Thus, x^* is weakly efficient if and only if, there does not exist $<h_j|\ j\epsilon I(x^*)>$, $<h_j^{(1)}|j\notin I(x^*)>$, $<h_j^{(2)}|j\notin I(x^*)>$ all non-negative and $g\epsilon\mathbb{R}_+^p$ such that:

$\sum_{j\in I(x^*)} A^j h_j + \sum_{j\notin I(x^*)} A^j h_j^{(1)} - \sum_{j\notin I(x^*)} A^j h_j^{(2)} = 0,$

$\sum_{j\in I(x^*)} C^j h_j + \sum_{j\notin I(x^*)} C^j h_j^{(1)} - \sum_{j\notin I(x^*)} C^j h_j^{(2)} - I^{(p)}g = E^{(p)}$

The reason validating the last 'p' equations is Note 1.

By Farkas' lemma, x^* is weakly efficient if and only if there exists $\xi\epsilon\mathbb{R}^m$, $\lambda\epsilon\mathbb{R}^p$ such that $\xi^TA^j + \lambda^TC^j \leq 0$ for all $j\epsilon\{1, \ldots, n\}$, $\xi^TA^j + \lambda^TC^j \leq 0$ for all $j\notin I(x^*)$, $-\lambda^T \leq 0$ and $\lambda^TE^{(p)} > 0$.

Thus, x^* is weakly efficient if and only if there exists $\xi\epsilon\mathbb{R}^m$, $\lambda\epsilon\mathbb{R}^p$ such that $\xi^TA^j + \lambda^TC^j \leq 0$ for all $j\epsilon I(x^*)$, $\xi^TA^j + \lambda^TC^j = 0$ for all $j\notin I(x^*)$ and $\lambda\epsilon\mathbb{R}_+^p\backslash\{0\}$.

Thus, if x^* is weakly efficient then there exists $\lambda\epsilon\mathbb{R}_+^p\backslash\{0\}$ such that for all $h\epsilon\mathbb{R}^n$ satisfying $h_j \geq 0$ for all $j\epsilon I(x^*)$, $Ah = 0$, it is the case that $\lambda^TCh \leq 0$.

For $h\epsilon\mathbb{R}^n$, $x^* + th \epsilon X(A,b)$ for all $t > 0$ sufficiently small if and only if [$h_j \geq 0$ for all $j\epsilon I(x^*)$ and $Ah = 0$].

Thus, if x^* is weakly efficient then there exists $\lambda\epsilon\mathbb{R}_+^p\backslash\{0\}$ such that $\lambda^TCx^* \geq \lambda^TCx$ for all $x\epsilon X(A, b)$.

Now suppose there exists $\lambda\epsilon\mathbb{R}_+^p\backslash\{0\}$ such that $\lambda^TCx^* \geq \lambda^TCx$ for all $x\epsilon X(A, b)$.

Towards a contradiction suppose there exists $x\epsilon X(A, b)$ such that $Cx >> Cx^*$.

Since $\lambda\epsilon\mathbb{R}_+^p\backslash\{0\}$, $Cx >> Cx^*$ implies $\lambda^TCx > \lambda^TCx^*$, contrary to our assumption.

Thus, x^* must be weakly efficient. Q.E.D.

Note 3: The interpretation of theorem 2 in our context is that x^* is efficient <u>if and only if</u> there exists a strictly positive array of probabilities over $\{1, \ldots, p\}$ with respect to which the decision maker maximizes the expected value of the objective function subject to the given linear constraints. ▌

As an immediate consequence of theorem 2 and duality theory of linear programming we have the following corollary.

Corollary 1 of Theorem 2: $x^*\epsilon X(A, b)$ is weakly efficient if and only if there exists $\lambda\epsilon\mathbb{R}_+^p\backslash\{0\}$ and $y^*\epsilon\mathbb{R}^m$ such that $y^{*T}A - \lambda^TC \geq 0$ and $(y^{*T}A - \lambda^TC)x^* = 0$.

4. Max-min Expected Value: Let $\Delta^{p-1} = \{\lambda\epsilon\mathbb{R}_+^p | \sum_{h=1}^p \lambda_h = 1\}$. Δ^{p-1} is the set of all probability vectors on $\{1, \ldots, p\}$.

Suppose that for each $h\epsilon\{1, \ldots, p\}$ there exists real numbers α_h, $\beta_h\epsilon [0, 1]$ satisfying $\beta_h \geq \alpha_h$ for all $h\epsilon\{1, \ldots, p\}$ and $\sum_{h=1}^p \beta_h \geq 1 \geq \sum_{h=1}^p \alpha_h$, such that the probability of the hth

objective function being realized belongs to the closed interval $[\alpha_h, \min\{\beta_h, 1\}]$. Let α denote the *p-vector* whose h^{th} coordinate is α_h, β denote the *p-vector* whose h^{th} coordinate is β_h and let $\Delta(\alpha, \beta) = \{\lambda \epsilon \Delta^{p-1} | \lambda_h \epsilon [\alpha_h, \beta_h]\}$.

The pay-off to the decision maker when it chooses $x \epsilon X(A, b)$ and nature chooses the randomization (mixed strategy) $\lambda \epsilon \Delta(\alpha, \beta)$ is $\lambda^T C x = x^T C^T \lambda$.

Here we assume that the decision-maker chooses an alternative that maximizes its minimum expected pay-off, where the minimization takes place across all possible imprecise randomizations.

Thus, the problem faced by the decision-maker is $\max_{x \epsilon X(A,b)} \min_{\lambda \epsilon \Delta(\alpha, \beta)} \lambda^T C x$.

$x^* \epsilon X(A, b)$ is a **max-min solution** if $x^* \epsilon \operatorname{argmax}_{x \epsilon X(A,b)} \{ \min_{\lambda \epsilon \Delta(\alpha, \beta)} \lambda^T C x \}$.

It is easy to see that if x^* solves the problem for the decision maker, then it must be weakly efficient.

It is well known that $\Delta(\alpha, \beta)$ is a polytope and hence is equal to the convex hull of its non-empty and finite set of extreme points. Let $\Delta^*(\alpha, \beta)$ denote the set of extreme points of $\Delta(\alpha, \beta)$, so that $\Delta(\alpha, \beta) = \{\sum_{\lambda \epsilon \Delta^*(\alpha, \beta)} \eta(\lambda)\lambda |\ \eta(\lambda) \geq 0$ for all $\lambda \epsilon \Delta^*(\alpha, \beta)$ and $\sum_{\lambda \epsilon \Delta^*(\alpha, \beta)} \eta(\lambda) = 1\}$.

It is also well known that X(A, b) is a polyhedron. Suppose X(A, b) is bounded. Then X(A, b) is a polytope and hence is equal to the convex hull of its non-empty and finite set of extreme points. Let $X^*(A, b)$ denote the set of extreme points of X(A, b), so that $X(A, b) = \{\sum_{x \epsilon X^*(A, b)} \xi(x) x |\ \xi(x) \geq 0$ for all $x \epsilon X^*(A, b)$ and $\sum_{x \epsilon X^*(A, b)} \xi(x) = 1\}$.

Note 4: While boundedness of X(A, b) is also sufficient for any linear programming problem of the form [Maximize $c^T x$, subject to $x \epsilon X(A, b)$] to have a solution, it is not necessary for the same purpose as the following problem illustrates: Maximize $x_1 + 2(x_2 - x_3)$, subject to $x_1 + x_2 - x_3 = 1$, $x_j \geq 0$, $j = 1, 2, 3$. In this case $\{(x_1, x_2, x_3) |\ x_1 + x_2 - x_3 = 1, x_j \geq 0, j = 1, 2, 3\}$ is unbounded, but the linear programming problem has a non-empty set of solutions $\{(0, x_2, 1 - x_2) |\ x_2 \epsilon [0, 1]\}$. Thus, although X(A, b) is not a polytope, the linear programming problem has a solution. ▌

Let ξ denote an array of the form $<\xi(x) | \xi(x) \geq 0$ for all $x \epsilon X^*(A, b)$ and $\sum_{x \epsilon X^*(A, b)} \xi(x) = 1>$ and η denote an array of the form $<\eta(\lambda) | \eta(\lambda) \geq 0$ for all $\lambda \epsilon \Delta^*(\alpha, \beta)$ and $\sum_{\lambda \epsilon \Delta^*(\alpha, \beta)} \eta(\lambda) = 1>$. Let Ψ be the set of all ξ's and H denote the set of all η's.

Thus, the problem faced by the decision maker is equivalent to the following problem:

$\max_{\xi\epsilon\Psi}\min_{\eta\epsilon H}(\sum_{\lambda\epsilon\Delta^*(\alpha)}(\sum_{\lambda\epsilon\Delta^*(\alpha)}\eta(\lambda)\lambda)^T C(\sum_{x\epsilon X^*(A,b)}\xi(x)x)$ or alternatively,

$\max_{\xi\epsilon\Psi}\min_{\eta\epsilon H}(\sum_{\lambda\epsilon\Delta^*(\alpha,\beta)}\sum_{x\epsilon X^*(A,b)}\xi(x)x)^T)C^T(\sum_{\lambda\epsilon\Delta^*(\alpha)}\eta(\lambda)\lambda)$.

We will say that $\xi^*\epsilon\Psi$ **solves** the problem above, if:$\xi^*\epsilon$

$\arg\max_{\xi\epsilon\Psi}\min_{\eta\epsilon H}(\sum_{\lambda\epsilon\Delta^*(\alpha)}(\sum_{\lambda\epsilon\Delta^*(\alpha,\beta)}\eta(\lambda)\lambda)^T C(\sum_{x\epsilon X^*(A,b)}\xi(x)x)$.

Important definition: For $(x, \lambda)\epsilon X^*(A, b)\times\Delta^*(\alpha, \beta)$, let $C(x, \lambda) = \lambda^T Cx = x^T C^T\lambda$.

The following result follows from the main result in Dantzig(1951).

Proposition 1: $\xi^*\epsilon\Psi$ solves the "equivalent" version of the decision maker's problem <u>if and only if</u> ξ^* solves the linear programming maximization problem: Maximize v, subject to $\sum_{x\epsilon X^*(A,b)}C(x,\lambda)\xi(x)$ - $v \geq 0$ for all $\lambda\epsilon\Delta^*(\alpha, \beta)$, $\sum_{x\epsilon X^*(A,b)}\xi(x) = 1$, $\xi(x) \geq 0$ for all $x\epsilon X^*(A, b)$, $v\epsilon\mathbb{R}$.

Note 5: Since $\Delta(\alpha, \beta)$ is a polytope with its set of extreme points being $\Delta^*(\alpha, \beta)$, $\min_{\lambda\epsilon\Delta^*(\alpha,\beta)}\sum_{x\epsilon X^*(A,b)}C(x,\lambda)\xi(x) = \min_{\lambda\epsilon\Delta(\alpha,\beta)}\sum_{x\epsilon X^*(A,b)}C(x,\lambda)\xi(x)$. Thus, $\max_{\xi\epsilon\Psi}\{\min_{\lambda\epsilon\Delta^*(\alpha,\beta)}\sum_{x\epsilon X^*(A,b)}C(x,\lambda)\xi(x) = \max_{\xi\epsilon\Psi}\{\min_{\lambda\epsilon\Delta(\alpha,\beta)}\sum_{x\epsilon X^*(A,b)}C(x,\lambda)\xi(x)\}$. This argument is used in the proof of proposition 1. ▌

An immediate consequence of proposition 1 is the following theorem.

Theorem 3: x^* is a max-min alternative <u>if and only if</u> there exists $\xi^*\epsilon\Psi$ satisfying the following properties:

(i) $x^* = \sum_{x\epsilon X^*(A,b)}\xi^*(x)x$.

<u>&</u>

(ii) ξ^* solves the linear programming maximization problem: Maximize v, subject to $\sum_{x\epsilon X^*(A,b)}C(x,\lambda)\xi(x)$ - $v \geq 0$ for all $\lambda\epsilon\Delta^*(\alpha, \beta)$, $\sum_{x\epsilon X^*(A,b)}\xi(x) = 1$, $\xi(x) \geq 0$ for all $x\epsilon X^*(A, b)$, $v\epsilon\mathbb{R}$.

Acknowledgment: I thank ChatGPT for pointing out errors in the proofs of theorems 1 and 2, errors in the formulation of the definition of max-min as well as an error in the example in note 3, in addition to other linguistic errors.

References

1. Dantzig, G. B. (1951): A Proof of the equivalence of the programming problem and the game problem. In: Activity Analysis of Production and Allocation, T. C. Koopmans (eds.). John Wiley & Sons, New York, Pages 330-335.

2. Ellsberg, D. (1961): Risk, Ambiguity and Savage Axioms. The Quarterly Journal of Economics, Volume 75, Number 4, Pages 643-649.

3. Gilboa, I. (2011): Making Better Decisions: Decision Theory in Practice. John Wiley Inc., Chichester.
4. Gilboa, I. and Schmeidler, D. (1989): Maxmin expected utility with a non-unique prior. Journal of Mathematical Economics, Volume 18, Pages 141-153.
5. Lahiri, S. (2026): Resolving Ambiguity in Multi-objective Linear Programming. DOI: https://doi.org/10.31224/8770
6. Luc, D. T. (2016): Multiobjective Linear Programming: An Introduction. Springer International Publishing.
7. Luptacik, M. (2010): Mathematical Optimization And Economic Analysis. Volume 36 of Springer Optimization and Its Applications. Springer, New York.