How to measure efficiency?

GameComm'2007

Arnaud Legrand, Corinne Touati

CNRS/INRIA, LIG laboratory

Oct. 22nd, 2007

Motivation

In the context of applied game theory in networking environments, a number of concepts have been proposed to measure both efficiency and optimality of resource allocations, the most famous certainly being the price of anarchy and the Jain index.

Yet, very few have tried to question and compare these measures one to another, in a general framework...

Outline

Outline

System:

- Consider an n-player game,
- Utility functions have values in $\mathbb{R}_+$,
- The utility set is U, a subset of $\mathbb{R}_+^n$.

General notations:

- $\mathcal{H}(\mathbb{R}_+^n)$ is the set of non-empty compact sets of $\mathbb{R}_+^n$,
- $\mathcal{C}(\mathbb{R}_+^n)$ is the set of non-empty compact and convex sets of $\mathbb{R}_+^n$.

We assume that $\mathcal{U}$ the set of all utility sets is either equal to $\mathcal{H}(\mathbb{R}_+^n)$ or $\mathcal{C}(\mathbb{R}_+^n)$. Any negative result regarding $\mathcal{C}(\mathbb{R}_+^n)$ also applies to $\mathcal{H}(\mathbb{R}_+^n)$.

Introduction: Two kinds of policy functions

Definition: Policy functions.

A (memoryless) policy function $\alpha : \mathcal{U} \to \mathbb{R}_+^n$ is a function such that $\forall U \in \mathcal{U}$, $\alpha(U) \in U$.
If $\mathcal{U} = \mathcal{H}(\mathbb{R}_+^n)$ (resp. $\mathcal{U} = \mathcal{C}(\mathbb{R}_+^n)$) then α is said to be a general policy function (resp. a convex policy function).

They do not necessarily optimize a specific function.
Example: Nash Equilibrium.

Definition: Policy functions.

A (memoryless) policy function $\alpha : \mathcal{U} \to \mathbb{R}_+^n$ is a function such that $\forall U \in \mathcal{U}$, $\alpha(U) \in U$.
If $\mathcal{U} = \mathcal{H}(\mathbb{R}_+^n)$ (resp. $\mathcal{U} = \mathcal{C}(\mathbb{R}_+^n)$) then α is said to be a general policy function (resp. a convex policy function).

They do not necessarily optimize a specific function.
Example: Nash Equilibrium.

Definition: "Index-optimizing" policy functions.

An index function f is a function from $\mathbb{R}_+^n$ to $\mathbb{R}_+$.
A policy function α is said to be f-optimizing if
$\forall U \in \mathcal{U}, f(\alpha(U)) = \sup_{u \in U} f(u)$.

They result on the optimization of a given function,
Examples: the Nash Bargaining Solution (proportional fairness), the social optimum, α-fairness...

- **Arithmetic mean**: $\sum_i u_i$.
- **Minimum**: $\min_i u_i$.
- **Maximum**: $\max_i u_i$.
- **Geometric Mean**: also called Nash Bargaining Solution or proportional fairness $\prod_i u_i$.
- **Harmonic Mean**: $1/(\sum_i 1/u_i)$.
- **Quasi-arithmetic Mean**: $f^{-1}(\frac{1}{n}\sum_{i=1}^{n} f(u_i))$ where f is a strictly monotone continuous function on $[0, +\infty]$. The particular case $f : x \to x^\delta$ is the α-fairness.
- **Jain**: $\frac{(\sum u_i)^2}{n\sum u_i^2}$.
- **Ordered Weighted Averaging**: $\sum_i w_i.u_{\sigma(i)}$ where σ is a permutation such that $u_{\sigma(1)} \leqslant u_{\sigma(2)} \leqslant \cdots \leqslant u_{\sigma(n)}$.

All these indexes are continuous, but not all strictly monotone.

Outline

- Pareto Optimality

Definition: Canonical order.

We also define the strict partial order $\ll$ on $\mathbb{R}_+^n$, namely the strict Pareto-superiority, by $u \ll v \Leftrightarrow \forall k : u_k < v_k$.

Definition: Pareto optimality.

A choice $u \in U$ is said to be Pareto optimal if it is maximal in U for the canonical partial order on $\mathbb{R}_+^n$.
A policy function is said to be Pareto-optimal if $\forall U \in \mathcal{U}, \alpha(U)$ is Pareto-optimal.

Nota: Even in systems that consists of independent elements, the Pareto optimality cannot be determined on each independent subsystem [Legrand, Touati, Infocom'07].

Qualitative Characterization: classical definitions

- Pareto Optimality
- Index-increasing

Definition: index-increasing.

A policy α is said to be f-increasing if $f \circ \alpha$ is monotone. Any f-optimizing policy is thus f-increasing.

Qualitative Characterization: classical definitions

- Pareto Optimality
- Index-increasing
- Braess-Paradoxes

Definition: Braess-paradox.

A policy function α is said to have Braess-paradoxes it there exists U_1 and U_2 such that

$$U_1 \subset U_2 \text{ and } \alpha(U_1) \gg \alpha(U_2)$$

A policy function such that there is no Braess-paradox is called Braess-paradox-free.

Pareto-optimality versus Index optimization

Theorem 1.

Let α be an f-optimizing policy. If f is strictly monotone then α is Pareto-optimal.

Pareto-optimality versus Index optimization

Theorem 1.

Let α be an f-optimizing policy. If f is strictly monotone then α is Pareto-optimal.

Theorem 2.

Let α be an f-optimizing policy. If α is Pareto-optimal then f is monotone.

Pareto-optimality versus Index optimization

Theorem 1.

Let α be an f-optimizing policy. If f is strictly monotone then α is Pareto-optimal.

Theorem 2.

Let α be an f-optimizing policy. If α is Pareto-optimal then f is monotone.

Nota: The Jain index is not monotone. Jain-optimizing policies are not necessarily Pareto optimal.

Examples: Min, $f : u \mapsto u_1$ are non strictly monotone indexes.

Continuity of allocations

Why is continuity important?

- In a dynamically changing systems, it ensures that a slight change in the resources would not significantly affect the allocation.
- It ensures that a slight error in the utility functions does not affect too much the allocation.

Continuity of allocations

Why is continuity important?

- In a dynamically changing systems, it ensures that a slight change in the resources would not significantly affect the allocation.
- It ensures that a slight error in the utility functions does not affect too much the allocation.

Definition: Hausdorff metric.

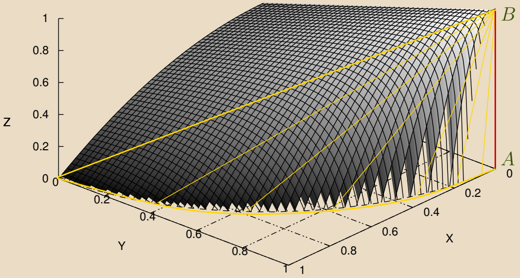
Let d be a metric function on $\mathbb{R}_+^n$.
The distance from x to compact B is $d(x, B) = \min\{d(x, y) | y \in B\}$.
The distance from the compact A to the compact B is $d(A, B) = \max\{d(x, B) | x \in A\}$.
The Hausdorff distance between two compacts A and B is then: $h(A, B) = \max(d(A, B), d(B, A))$

Nota: $(\mathcal{H}(\mathbb{R}_+^n), h)$ and $(\mathcal{C}(\mathbb{R}_+^n), h)$ are complete metric spaces.

Theorem 3.

- The Pareto set of a convex utility set is not necessarily compact.
- The function $\overline{\mathcal{P}}$ from $\mathcal{C}(\mathbb{R}_+^n)$ to $\mathcal{H}(\mathbb{R}_+^n)$ that associates to U the closure of its Pareto set is not continuous.

Proof.



Convex set whose Pareto set is not closed. The segment $[A, B[$ does not belong to the Pareto set.

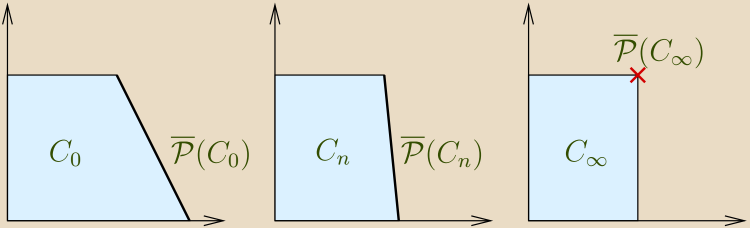
□

Continuity of allocations

Theorem 3.

- The Pareto set of a convex utility set is not necessarily compact.
- The function $\overline{\mathcal{P}}$ from $\mathcal{C}(\mathbb{R}_+^n)$ to $\mathcal{H}(\mathbb{R}_+^n)$ that associates to U the closure of its Pareto set is not continuous.

Proof.



□

Continuity of allocations

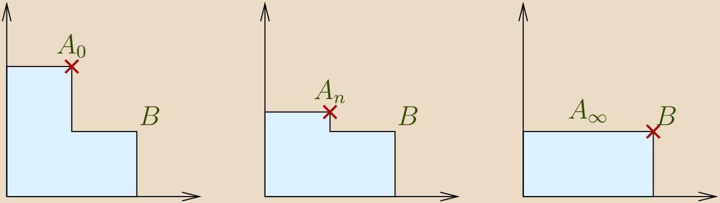
Theorem 3.

- The Pareto set of a convex utility set is not necessarily compact.
- The function $\overline{\mathcal{P}}$ from $\mathcal{C}(\mathbb{R}_+^n)$ to $\mathcal{H}(\mathbb{R}_+^n)$ that associates to U the closure of its Pareto set is not continuous.

Theorem 4.

Let α be a general Pareto-optimal policy function. α is not continuous.

Proof.



□

Continuity of allocations

Theorem 3.

- The Pareto set of a convex utility set is not necessarily compact.
- The function $\overline{\mathcal{P}}$ from $\mathcal{C}(\mathbb{R}_+^n)$ to $\mathcal{H}(\mathbb{R}_+^n)$ that associates to U the closure of its Pareto set is not continuous.

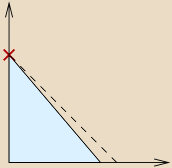
Theorem 4.

Let α be a general Pareto-optimal policy function. α is not continuous.

Theorem 5.

There exists continuous and non-continuous convex Pareto-optimal policy functions.

Proof.



Sum-optimizing is discontinuous, but the geometric mean-optimizing policy is continuous. □

Continuity of allocations

Theorem 3.

- The Pareto set of a convex utility set is not necessarily compact.
- The function $\overline{\mathcal{P}}$ from $\mathcal{C}(\mathbb{R}_+^n)$ to $\mathcal{H}(\mathbb{R}_+^n)$ that associates to U the closure of its Pareto set is not continuous.

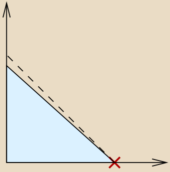
Theorem 4.

Let α be a general Pareto-optimal policy function. α is not continuous.

Theorem 5.

There exists continuous and non-continuous convex Pareto-optimal policy functions.

Proof.



Sum-optimizing is discontinuous, but the geometric mean-optimizing policy is continuous. □

Theorem 6.

Let f and g be two monotone index functions. A g-optimizing policy α_g is f-increasing if and only if α_g is f-optimizing.

In other words, a policy optimizing an index f is always non-monotone for a distinct index g.
This explains why allocations that are efficient (optimizing the arithmetic mean) cannot (in general) also be fair (optimizing the geometric mean).

Monotonicity

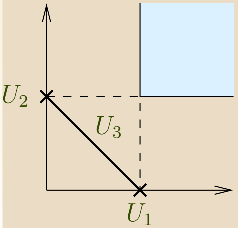
Theorem 6.

Let f and g be two monotone index functions. A g-optimizing policy α_g is f-increasing if and only if α_g is f-optimizing.

Theorem 7.

Even if convex, policy functions cannot be monotone.

Proof.



□

This infers that even in Braess-free systems, an increase in the resource can be detrimental to some users.

Qualitative Characterizations: summary

- Indexes should be strictly monotone to ensure Pareto-Optimality.

Qualitative Characterizations: summary

- Indexes should be strictly monotone to ensure Pareto-Optimality.
- Continuity (of allocations) is only possible when considering convex utility sets.

Qualitative Characterizations: summary

- Indexes should be strictly monotone to ensure Pareto-Optimality.
- Continuity (of allocations) is only possible when considering convex utility sets.
- A policy optimizing a given index f leads to erratic values of an other index g when growing utility sets (unless f and g induce the same optimization).

Qualitative Characterizations: summary

- Indexes should be strictly monotone to ensure Pareto-Optimality.
- Continuity (of allocations) is only possible when considering convex utility sets.
- A policy optimizing a given index f leads to erratic values of an other index g when growing utility sets (unless f and g induce the same optimization).
- It is impossible to ensure that the growth of the utility set does not induce the decrease of the utility of some player (even when restricting to convex utility sets).

Qualitative Characterizations: summary

- Indexes should be strictly monotone to ensure Pareto-Optimality.
- Continuity (of allocations) is only possible when considering convex utility sets.
- A policy optimizing a given index f leads to erratic values of an other index g when growing utility sets (unless f and g induce the same optimization).
- It is impossible to ensure that the growth of the utility set does not induce the decrease of the utility of some player (even when restricting to convex utility sets).

Even though being Braess-paradox-free does not lead to bad properties, it does not give any information on how efficient the policies are.

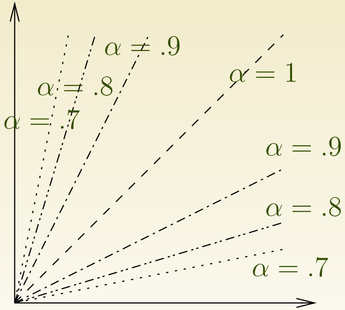
Outline

Jain index

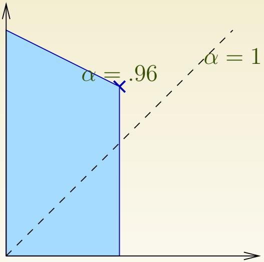
The Jain efficiency measure (or Jain index) of a choice u is defined as $\frac{(\sum u_i)^2}{n \sum u_i^2}$. The Jain index is thus the ratio of the first to the second moment of the choice u. Hence, it is considered as a good measure of a choice fairness (as defined by max-min fairness).

Jain index

The Jain efficiency measure (or Jain index) of a choice u is defined as $\frac{(\sum u_i)^2}{n \sum u_i^2}$. The Jain index is thus the ratio of the first to the second moment of the choice u. Hence, it is considered as a good measure of a choice fairness (as defined by max-min fairness).



(a) The Jain-index is not monotone: hence optimal solutions for the index may not be Pareto-optimal.



(b) The max-min fair allocation may have a sub-optimal Jain index!

Price of anarchy

For a given index f, let us consider $\alpha^{(f)}$ an f-optimizing policy function. We define the inefficiency $I_f(\beta, U)$ of the allocation $\beta(U)$ for f as

$$I_f(\beta, U) = \frac{f(\alpha^{(f)}(U))}{f(\beta(U))} = \max_{u \in U} \frac{f(u)}{f(\beta(U))} \geqslant 1$$

Price of anarchy

For a given index f, let us consider $\alpha^{(f)}$ an f-optimizing policy function. We define the inefficiency $I_f(\beta, U)$ of the allocation $\beta(U)$ for f as

$$I_f(\beta, U) = \frac{f(\alpha^{(f)}(U))}{f(\beta(U))} = \max_{u \in U} \frac{f(u)}{f(\beta(U))} \geqslant 1$$

Papadimitriou focuses on the arithmetic mean Σ defined by

$$\Sigma(u_1, \ldots, u_k) = \sum_{k=1}^{K} u_k$$

The price of anarchy ϕ_Σ is thus defined as the largest inefficiency:

$$\phi_\Sigma(\beta) = \sup_{U \in \mathcal{U}} I_f(\beta, U) = \sup_{U \in \mathcal{U}} \frac{\sum_k \alpha^{(\Sigma)}(U)_k}{\sum_k \beta(U)_k}$$

Price of anarchy

For a given index f, let us consider $\alpha^{(f)}$ an f-optimizing policy function. We define the inefficiency $I_f(\beta, U)$ of the allocation $\beta(U)$ for f as

$$I_f(\beta, U) = \frac{f(\alpha^{(f)}(U))}{f(\beta(U))} = \max_{u \in U} \frac{f(u)}{f(\beta(U))} \geqslant 1$$

Papadimitriou focuses on the arithmetic mean Σ defined by

$$\Sigma(u_1, \ldots, u_k) = \sum_{k=1}^{K} u_k$$

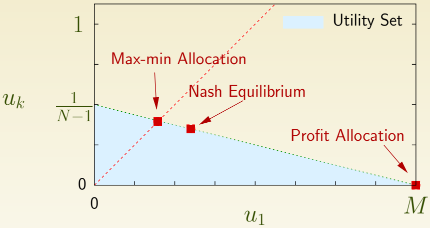
The price of anarchy ϕ_Σ is thus defined as the largest inefficiency:

$$\phi_\Sigma(\beta) = \sup_{U \in \mathcal{U}} I_f(\beta, U) = \sup_{U \in \mathcal{U}} \frac{\sum_k \alpha^{(\Sigma)}(U)_k}{\sum_k \beta(U)_k}$$

In other words, $\phi_\Sigma(\beta)$ is the approximation ratio of β for the objective function Σ.

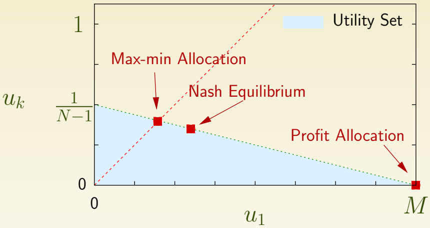
Price of anarchy: does it really reflects inefficiencies?

Consider the utility set $S_{M,N} = \{u \in \mathbb{R}_+^N | \frac{u_1}{M} + \sum_{k=1}^N u_k \leqslant 1\}$. As the roles of the u_k, $k \geqslant 2$ are symmetric, we can freely assume that $u_2 = \cdots = u_N$ for index-optimizing policies ([Legrand et al, Infocom'07]).



Utility set and allocations for $S_{M,N}$ ($N = 3$,$M = 2$), with $u_2 = \cdots = u_N$.

Consider the utility set $S_{M,N} = \{u \in \mathbb{R}_+^N | \frac{u_1}{M} + \sum_{k=1}^{N} u_k \leqslant 1\}$. As the roles of the u_k, $k \geqslant 2$ are symmetric, we can freely assume that $u_2 = \cdots = u_N$ for index-optimizing policies ([Legrand et al, Infocom'07]).

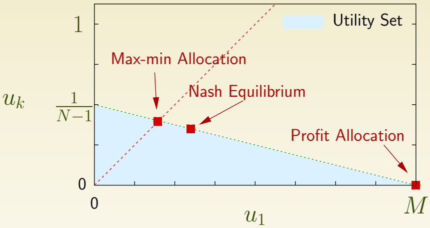


Utility set and allocations for $S_{M,N}$ ($N = 3$, $M = 2$), with $u_2 = \cdots = u_N$.

$$I_\Sigma(\alpha^{\text{NBS}}, S_{M,N}) \xrightarrow[M\to\infty]{} N$$

$$I_\Sigma(\alpha^{\text{Max-Min}}, S_{M,N}) \sim_{M\to\infty} M$$

Consider the utility set $S_{M,N} = \{u \in \mathbb{R}_+^N | \frac{u_1}{M} + \sum_{k=1}^N u_k \leqslant 1\}$. As the roles of the u_k, $k \geqslant 2$ are symmetric, we can freely assume that $u_2 = \cdots = u_N$ for index-optimizing policies ([Legrand et al, Infocom'07]).



Utility set and allocations for $S_{M,N}$ $(N = 3, M = 2)$, with $u_2 = \cdots = u_N$.

$$I_\Sigma(\alpha^{\text{NBS}}, S_{M,N}) \xrightarrow[M\to\infty]{} N$$

$$I_\Sigma(\alpha^{\text{Max-Min}}, S_{M,N}) \sim_{M\to\infty} M$$

These are due to the fact that a policy optimizing an index f is always non-monotone for a distinct index g.

$\rightsquigarrow$ Pareto inefficiency should be measured as the distance to the Pareto border and not to a specific point.

Selfishness Degradation Factor

Definition: Pareto-ratio.

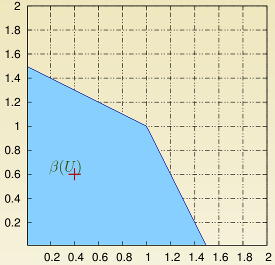
α is strictly superior to β iff $\varrho(\alpha, \beta) = \min_k \frac{\alpha_k}{\beta_k} > 1$ [Kameda, Networking'04].

Definition: Selfishness Degradation Factor (SDF).

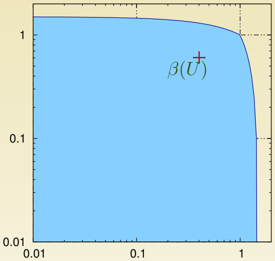
The Selfishness Degradation Factor (SDF) is defined by [Legrand, Touati, Infocom'07]:

$$I_{SDF}(\beta, U) = \max_{u \in U} \varrho(u, \beta(U)) = \max_{u \in U} \min_k \frac{u_k}{\beta(U)_k}$$

Therefore $\beta(U)$ is Pareto inefficient as soon as $I_{SDF}(\beta, U) > 1$ and the larger $I_{SDF}(\beta, U)$, the more inefficient the allocation.



U and $\beta(U)$ in the original space

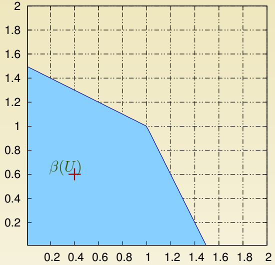


U and $\beta(U)$ in the $\log$-space

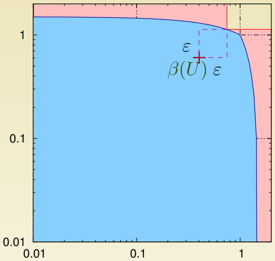
Lemma 1.

Let us define $\llcorner = \{x \in \mathbb{R}^n | \exists k : x_k \leqslant 0\}$. We denote by $a \boxplus \llcorner = \{x \in \mathbb{R}^n | \exists k : x_k \leqslant a_k\}$. Then:

$$\log(I_{SDF}(\beta, U)) \leqslant \varepsilon \Leftrightarrow \log(U) \subseteq (\log(\beta(U)) + \varepsilon) \boxplus \llcorner$$



U and $\beta(U)$ in the original space

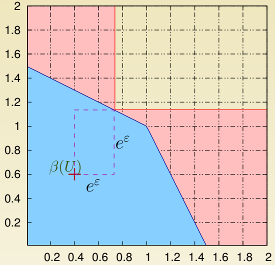


U and $\beta(U)$ in the $\log$-space

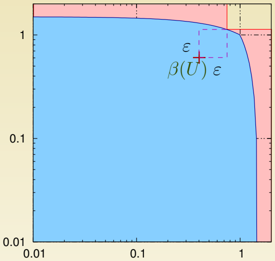
Lemma 1.

Let us define $\llcorner = \{x \in \mathbb{R}^n | \exists k : x_k \leqslant 0\}$. We denote by $a \boxplus \llcorner = \{x \in \mathbb{R}^n | \exists k : x_k \leqslant a_k\}$. Then:

$$\log(I_{SDF}(\beta, U)) \leqslant \varepsilon \Leftrightarrow \log(U) \subseteq (\log(\beta(U)) + \varepsilon) \boxplus \llcorner$$



U and $\beta(U)$ in the original space



U and $\beta(U)$ in the $\log$-space

Lemma 1.

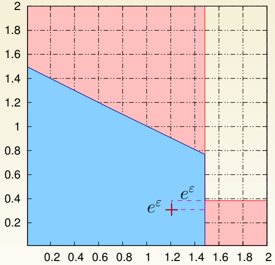
Let us define $\mathbb{L} = \{x \in \mathbb{R}^n | \exists k : x_k \leqslant 0\}$. We denote by $a \boxplus \mathbb{L} = \{x \in \mathbb{R}^n | \exists k : x_k \leqslant a_k\}$. Then:

$$\log(I_{SDF}(\beta, U)) \leqslant \varepsilon \Leftrightarrow \log(U) \subseteq (\log(\beta(U)) + \varepsilon) \boxplus \mathbb{L}$$

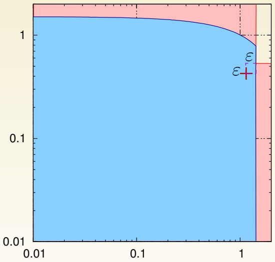
Selfishness Degradation Factor: graphical interpretation

- This inefficiency seems to measure how much $\beta(U)$ should be increased so that it is not dominated by any other points in U.
- $\log(I_{SDF}(\beta, U))$ "somehow" measures the distance in the log-space from $\beta(U)$ to the Pareto set. However, this definition holds only because of the very specific shape of the set $\mathcal{U}$ used in this example.

- This inefficiency seems to measure how much $\beta(U)$ should be increased so that it is not dominated by any other points in U.
- $\log(I_{SDF}(\beta, U))$ "somehow" measures the distance in the log-space from $\beta(U)$ to the Pareto set. However, this definition holds only because of the very specific shape of the set $\mathcal{U}$ used in this example.



U and $\beta(U)$ in the original space



U and $\beta(U)$ in the $\log$-space

Topological Point of View: reformulation of the SDF

- The SDF can be properly defined only when referring to the whole Pareto set.
- The distance from $\beta(U)$ to the closure of the Pareto set $\overline{\mathcal{P}}(U)$ in the log-space is equal to:

$$d_\infty(\log(\beta(U), \log(\overline{\mathcal{P}}(U))) = \min_{u \in \overline{\mathcal{P}}(u)} \max_k |\log(\beta(U)_k) - \log(u_k)|$$

Therefore, we can define

$$\begin{aligned} \widetilde{I}_\infty(\beta, U) &= \exp(d_\infty(\log(\beta(U), \log(\overline{\mathcal{P}}(U))) \\ &= \min_{u \in \overline{\mathcal{P}}(u)} \max_k \max\left(\frac{\beta(U)_k}{u_k}, \frac{u_k}{\beta(U)_k}\right) \end{aligned} \quad (1)$$

SDF and ε-approximation

Definition: ε-approximation.

[Papadimitriou, Yannakakis, FOCS'00] defines an ε-approximation of $\overline{\mathcal{P}}(U)$ as a set of points S such that for all $u \in U$ there exists some $s \in S$ such that $\forall k : u_k \leqslant (1+\varepsilon)s_k$.

Definition: Expansion.

$$X \oplus a = \{y | d(x,y) \leqslant a, \text{ for some } x \in X\}$$

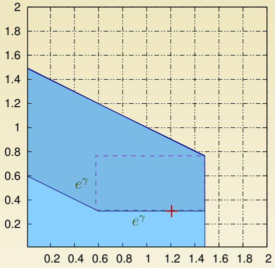
This definition can be easily expanded as:

$$\begin{aligned} X \otimes a &= \exp(\log(X) \oplus \log(a)) \\ &= \{y | \exp(d(\log(x), \log(y)) \leqslant a \text{ for some } x \in X\} \end{aligned}$$

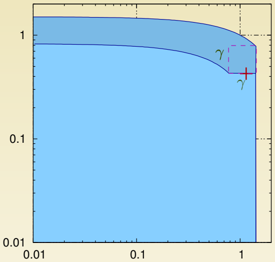
Theorem 8.

$S \subseteq U$ is an ε-approximation of $\overline{\mathcal{P}}(U)$ iff $\overline{\mathcal{P}}(U) \subseteq S \otimes \exp(\varepsilon)$.

SDF and ε-approximation

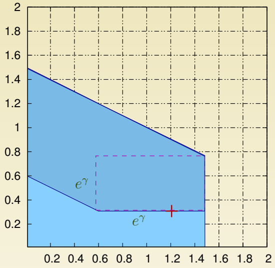


U and $\beta(U)$ in the original space

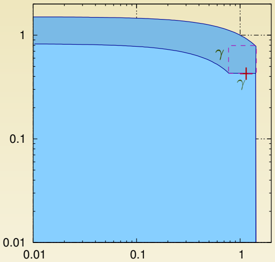


U and $\beta(U)$ in the $\log$-space

SDF and ε-approximation



U and $\beta(U)$ in the original space



U and $\beta(U)$ in the $\log$-space

Lemma 2.

$\widetilde{I}_\infty(\beta, U) \leqslant \exp(\varepsilon) \Leftrightarrow \beta(U) \in \overline{\mathcal{P}}(U) \otimes \exp(\varepsilon)$.

In other words, $\widetilde{I}_\infty(\beta, U) \leqslant \exp(\varepsilon)$ iff $\beta(U)$ is no farther than ε from $\overline{\mathcal{P}}(U)$ in the log space.

SDF: summary

- When comparing the definitions of I_Σ, I_{SDF} and $\widetilde{I}_\infty$, the latest may seem harder to compute as it relies on $\mathcal{P}(U)$.
- However, what we are interested in is measuring the distance to the Pareto set and no index-based inefficiency measure can reflect this distance. They can only reflect a particular property of the allocation such as fairness.
- Note that in mono-criteria situations, it is natural to compare a solution to an intractable optimal solution, generally using approximations or lower bounds. Therefore, similar approaches should be used in multi-criteria settings to compute $\widetilde{I}_\infty$. This inefficiency measure is thus a natural extension of the classical mono-criteria performance ratio.
- The previous definition should thus be used in the general case, even though in a some particular situations, the original SDF definition is sufficient.

Outline

Conclusion

We have addressed the question of how to properly measure efficiency of allocations, may they be obtained as the result of some index-function optimization or some general policy.

- Monotonicity is the link between index-optimization and Pareto optimality.
- When utilities are continuous with the system's resources, solution allocations can be continuous in the resources only when the utility sets are convex.
- Even with Braess-free allocations, there always exists instances where resource increase is detrimental to at least one user.
- A policy optimizing a given index leads to erratic values for another index when utility sets grow.
- Both the Jain index and the price of anarchy have flaws as measures of the inefficiency of an equilibria.
- A correct general inefficiency measure can be defined based on the log space as the distance of a point to the Pareto border.