MODELAGEM COMPUTACIONAL
PROGRAMA DE PÓS-GRADUAÇÃO

Otimização Estocástica

Princípios da Otimização Multi-objetivo

Gustavo Barbosa Libotte

27 a 29 de setembro de 2022

Visão geral

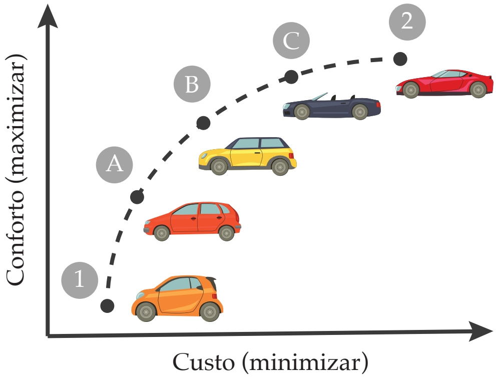
Otimização

A otimização refere-se a **encontrar uma ou mais soluções viáveis** que correspondam a valores extremos de um ou mais objetivos. A necessidade de encontrar tais soluções ótimas em um problema vem principalmente do propósito extremo de projetar uma solução para o custo mínimo possível de fabricação, ou para a máxima confiabilidade possível, dentre outros.

- Quando um problema de otimização modelando um sistema físico envolve apenas uma função objetivo, a tarefa de encontrar a solução ótima é chamada de **otimização de objetivo único**.
- Existem algoritmos de otimização de objetivo único que funcionam usando técnicas de pesquisa baseadas em **gradiente** (técnicas clássicas).
- Além dos princípios de busca determinísticos envolvidos em um algoritmo, existem também princípios de busca **estocásticos**.

- Quando um problema de otimização envolve mais de uma função objetivo, a tarefa de encontrar uma ou mais soluções ótimas é conhecida como **otimização multi-objetivo**.
- Como a otimização multi-objetivo envolve vários objetivos, é **intuitivo** pensar que a otimização de objetivo único é um **caso degenerado** de otimização multi-objetivo.
- A maioria dos problemas de busca e otimização do **mundo real** naturalmente envolve múltiplos objetivos.
- O princípio extremista **não pode** ser aplicado a apenas um objetivo, quando os demais objetivos também são importantes. Diferentes soluções podem produzir ***trade-offs*** (cenários conflitantes) entre diferentes objetivos. Uma solução extrema (no melhor sentido) em relação a um objetivo requer um **compromisso** em outros objetivos.

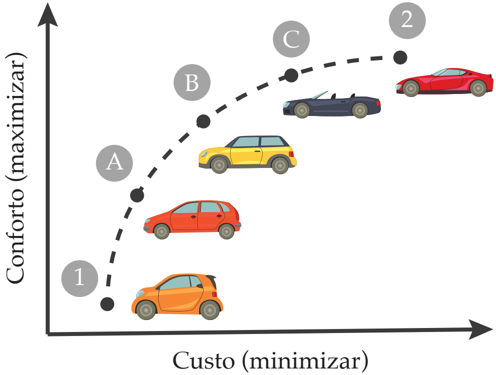
- Consideremos a tomada de decisão envolvida na **compra de um automóvel**.



- Os carros estão disponíveis a preços que variam de alguns milhares a algumas centenas de milhares de reais. Tomemos dois carros hipotéticos **extremos**, ou seja, um custando cerca de 40 mil reais (solução 1) e outro custando cerca de 300 mil reais (solução 2).
- Se o custo é o **único objetivo** deste processo de tomada de decisão, a escolha ótima seria a solução 1.
- Se este fosse o único objetivo para todos os compradores, teríamos visto apenas um tipo de carro (solução 1) na estrada.

VISÃO GERAL

- Felizmente, esse processo de tomada de decisão **não é de um objetivo único**. Salvo algumas exceções, espera-se que um carro barato seja menos confortável.



- Para os compradores ricos para quem o **conforto é o único objetivo** dessa tomada de decisão, a escolha é a solução 2 (com um nível de conforto máximo).
- Entre estas duas soluções extremas, existem muitas outras soluções **intermediárias**, onde existe um **compromisso entre custo e conforto**.
- Assim, entre quaisquer duas soluções, uma é **melhor** em termos de um objetivo, mas essa melhoria vem apenas de um **sacrifício** no outro objetivo.

Visão geral

- Para dois objetivos conflitantes, cada objetivo corresponde a uma **solução ótima diferente**.
- No problema hipotético de tomada de decisão de compra de um carro, as **soluções 1 e 2** são essas soluções ótimas.
- Se um comprador estiver disposto a **sacrificar o custo** até certo ponto da solução 1, provavelmente poderá encontrar outro carro com um nível de conforto melhor do que esta solução.
- Idealmente, a extensão do **sacrifício** no custo está relacionada ao **ganho** em conforto.
- Assim, podemos visualizar um conjunto de soluções ótimas onde um **ganho** em um objetivo exige um **sacrifício** no outro objetivo.

Solução ótima

Agora vem a grande questão: dentre todas as soluções, pode-se dizer qual é a **melhor solução** em relação a ambos os objetivos?

Visão geral

- A ironia é que **nenhuma** dessas soluções conflitantes é a melhor em relação a ambos os objetivos.
- A razão está no fato de que nenhuma solução deste conjunto faz com que **ambos os objetivos** (custo e conforto) pareçam melhores do que qualquer outra solução do conjunto.
- Assim, em problemas com mais de um objetivo conflitante, **não existe uma única solução ótima**.
- Sem qualquer **informação adicional**, nenhuma solução do conjunto de soluções ótimas pode ser considerada melhor do que qualquer outra.
- Esta é a **diferença fundamental** entre uma tarefa de otimização de objetivo único e multi-objetivo.
- Exceto na otimização multimodal de objetivo único, a solução importante em uma otimização de objetivo único é a **única solução ótima**, enquanto que na otimização multi-objetivo, **várias soluções ótimas** decorrentes de *trade-offs* entre objetivos conflitantes são importantes.

Visão geral

- Embora a diferença fundamental entre essas duas otimizações esteja na **cardinalidade no conjunto ótimo**, do ponto de vista prático um usuário precisa de apenas uma solução.
- Qual dessas soluções ótimas deve ser escolhida? Vamos tentar responder a esta pergunta para o caso do problema da **compra de carros**.
- Conhecendo o número de soluções que existem no mercado com diferentes *trade-offs* entre **custo e conforto**, qual carro comprar?
- Envolve muitas **outras considerações**, como o
  - recursos financeiros disponíveis para comprar o carro;
  - distância a ser percorrida a cada dia;
  - número de passageiros que viajam no carro;
  - consumo e custo de combustível;
  - condições da estrada onde o carro deve ser conduzido principalmente;
  - saúde física dos passageiros.

Visão geral

- Muitas vezes, essas informações de nível superior **não são** técnicas, qualitativas e baseadas na experiência.
- Pode-se avaliar os **prós e contras** de cada uma dessas soluções com base em todas essas considerações não técnicas e qualitativas, mas ainda importantes, e compará-las para fazer uma escolha.
- Assim, em uma otimização multi-objetivo, idealmente o esforço deve ser feito em encontrar o conjunto de soluções ótimas de *trade-off* considerando **todos os objetivos como importantes**.
- Depois que um conjunto de tais soluções é encontrado, um usuário pode usar **considerações qualitativas de nível superior** para fazer uma escolha.
- Esta tarefa se chama **pós-processamento de soluções**.
- Chamamos esse procedimento de otimização multi-objetivo **baseada em preferências**.

Problema de otimização multi-objetivo

- Problemas de otimização multi-objetivo são da forma:

$$\text{opt}\,\{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})\}$$
$$\text{Sujeito a } \mathbf{x} \in S$$

  onde temos $m \geq 2$ funções objetivo $f_i : \mathbb{R}^n \to \mathbb{R}$.

- Na otimização multi-objetivo, alguns dos objetivos devem ser **minimizados** e outros **maximizados**:
  - sejam $m_1 \geq 0$ e $m_2 \geq 0$, com $m = m_1 + m_2$, e sejam ordenados os objetivos f_i tal que os primeiros m_1 objetivos, para $i = 1, \ldots, m_1$, devem ser minimizados;
  - os próximos m_2 objetivos, para $i = m_1 + 1, \ldots, m_1 + m_2$, devem ser maximizados.

- Denotamos o vetor de **funções objetivo** por $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))^\top$.

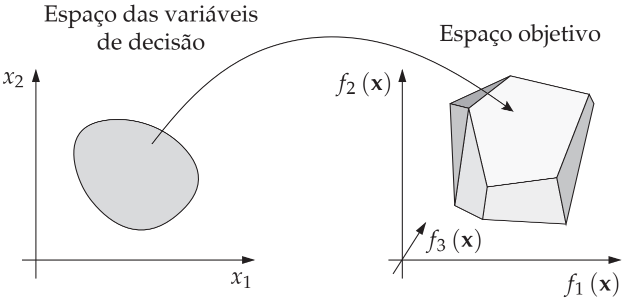
- Os **vetores de decisão** $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$ pertencem à **região viável** S, que é um subconjunto do **espaço das variáveis de decisão** $\mathbb{R}^n$.

- Ainda não fixamos a forma das **funções de restrição** que formam S, mas nos referimos a S em geral.

Problema de otimização multi-objetivo

- Denotamos a imagem da região viável por $Z = \mathbf{f}(S)$ e a chamamos de **região objetivo viável**. É um subconjunto do **espaço objetivo** $\mathbb{R}^m$.

- Os elementos de Z são chamados de vetores objetivos ou vetores de critério e denotados por $\mathbf{f}(\mathbf{x})$ ou $\mathbf{z} = (z_1, z_2, \ldots, z_m)^T$, onde $z_i = f_i(\mathbf{x})$ para todos os $i = 1, \ldots, m$ são valores objetivos ou valores de critério.



Problema de otimização multi-objetivo

- Em problemas de otimização de objetivo único, o foco principal está no **espaço das variáveis de decisão**.
- No contexto multi-objetivo, muitas vezes estamos mais interessados no **espaço objetivo**, que geralmente é de uma dimensão menor do que o espaço das variáveis de decisão.
- Por causa da contradição e possível incomensurabilidade das funções objetivo, **não é possível** encontrar uma única solução que seja ótima para todos os objetivos simultaneamente.
- Portanto, não há ordenação natural no espaço objetivo porque ele é apenas **parcialmente ordenado**.
- **Então, como escolher soluções que resolvam o problema de otimização multi-objetivo?**
- Vamos explorar mais a fundo esta questão. Mas antes, vamos falar um pouco sobre **relações binárias**.

Relação binária em um conjunto S

Uma relação binária em um conjunto S é um subconjunto de $S \times S$ (um conjunto de pares ordenados de elementos de S).

- **Exemplo:** Seja $S = \{1, 2, 4\}$. No conjunto

$$S \times S = \{(1,1), (1,2), (1,4), (2,1), (2,2), (2,4), (4,1), (4,2), (4,4)\},$$

pode-se definir uma relação ρ por $x \ \rho \ y$ se e somente se $x = y/2$, abreviada como $x \ \rho \ y \leftrightarrow x = y/2$. Assim, $(1,2)$ e $(2,4)$ satisfazem ρ. A mesma relação ρ poderia ser definida dizendo-se que $\{(1,2), (2,4)\}$ é o conjunto dos pares ordenados que satisfazem ρ.

- Agora que sabemos que uma relação binária ρ é um **subconjunto**, vemos que

$$x \ \rho \ y \leftrightarrow (x, y) \in \rho$$

Em geral, uma relação binária é definida por uma **descrição da relação**, em vez da lista dos pares ordenados.

- **Exemplo:** Seja $S = \{1,2\}$. Então, $S \times S = \{(1,1),(1,2),(2,1),(2,2)\}$. Seja ρ a relação em S dada pela descrição $x \; \rho \; y \leftrightarrow x + y$ é ímpar. Então $(1,2) \in \rho$ e $(2,1) \in \rho$. O par ordenado $(1,1) \notin \rho$ porque $1+1$ não é ímpar. Analogamente, $(2,2) \notin \rho$.

- Uma relação binária em um conjunto S pode ter determinadas **propriedades**.

- Por exemplo, a relação ρ de igualdade em S, $x \; \rho \; y \leftrightarrow x = y$, tem três propriedades:
  1. para qualquer $x \in S, x = x$, ou seja, $(x,x) \in \rho$;
  2. quaisquer que sejam $x,y \in S$, se $x = y$, então $y = x$, ou seja, $(x,y) \in \rho \to (y,x) \in \rho$;
  3. quaisquer que sejam $x,y,z \in S$, se $x = y$ e $y = z$, então $x = z$, ou seja, $[(x,y) \in \rho$ e $(y,z) \in \rho] \to (x,z) \in \rho$.

Reflexividade, simetria e transitividade em palavras

- Reflexividade: todo x está relacionado consigo mesmo.
- Simetria: se x estiver relacionado com y, então y estará relacionado com x.
- Transitividade: se x estiver relacionado com y e y estiver relacionado com z, então x estará relacionado com z.

- **Mais formalmente:** seja ρ uma relação binária em um conjunto S. Então,
  - Dado que $(\forall x)(x \in S \to (x,x) \in \rho)$, a relação ρ é dita ser **reflexiva**.
  - Dado que $(\forall x)(\forall y)(x \in S \wedge y \in S \wedge (x,y) \in \rho \to (y,x) \in \rho)$, a relação ρ é dita ser **simétrica**.
  - Dado que $(\forall x)(\forall y)(\forall z)(x \in S \wedge y \in S \wedge (x,y) \in \rho \wedge (y,z) \in \rho \to (x,z) \in \rho)$, a relação ρ é dita ser **transitiva**.
- **Exemplo:** considere a relação $\leq$ no conjunto $\mathbb{N}$.
  - Essa relação é **reflexiva** porque, para qualquer inteiro não negativo x, $x \leq x$.
  - Ela é, também, **transitiva** pois quaisquer que sejam os inteiros não negativos x, y, z, se $x \leq y$ e $y \leq z$, então $x \leq z$.
  - No entanto, $\leq$ **não é simétrica**; $3 \leq 4$ não implica que $4 \leq 3$. De fato, quaisquer que sejam $x, y \in \mathbb{N}$, se $x \leq y$ e $y \leq x$, então $x = y$.
  - Essa característica é descrita dizendo-se que $\leq$ é **antissimétrica**.
- **Definição:** Seja ρ uma relação binária em um conjunto S. Dizer que ρ é **antissimétrica** significa

$$(\forall x)(\forall y)(x \in S \wedge y \in S \wedge (x,y) \in \rho \wedge (y,x) \in \rho \to x = y)$$

Ordem parcial

Uma relação binária em um conjunto S que seja reflexiva, antissimétrica e transitiva é chamada de uma **ordem parcial** em S.

- Vejamos alguns **exemplos** de ordens parciais:
  - Em $\mathbb{N}$, $x \rho y \leftrightarrow x \leq y$;
  - Em Z^+, $x \rho y \leftrightarrow x$ divide y;
  - Em $\{0,1\}$, $x \rho y \leftrightarrow x = y^2$.
- Se ρ for uma ordem parcial em S, então o par ordenado (S, ρ) será chamado de um **conjunto parcialmente ordenado**.
- Denotaremos um conjunto parcialmente ordenado arbitrário por $(S, \preccurlyeq)$; em qualquer caso particular, $\preccurlyeq$ tem algum significado preciso, como "menor ou igual a" ou "é um subconjunto de".
- O símbolo para uma ordem parcial genérica, $\preccurlyeq$, é projetado para parecer com o símbolo de desigualdade $\leq$, que, como acabamos de observar, é uma ordem parcial no conjunto $\mathbb{N}$.

- Implícita em qualquer processo de tomada de decisão está a necessidade de construir, direta ou indiretamente, a **ordem de preferência**, para que as alternativas possam ser classificadas e a escolha final possa ser selecionada.
- Para alguns problemas de tomada de decisão, isso pode ser facilmente realizado.
- Por exemplo, pode ser apropriado tomar uma decisão com base em uma regra de minimização de custos, em que a ordem de preferência seja adequadamente representada pela **ordem natural dos números reais** ($\leq$ ou $\geq$).
- Em algumas situações, no entanto, pode ser desejável explorar a estrutura de **preferência do tomador de decisão** de alguma forma direta e tentar construir algum tipo de ordem de preferência diretamente.
- Devemos definir uma **relação de ordem** que pode ser usada para classificar alternativas de uma maneira que reflita a preferência do tomador de decisão.

- A ordem de preferência é uma **relação binária**.
- Usaremos o símbolo $\succ$ para denotar a ordem de preferência estrita "**é preferível a**"; o símbolo $\sim$ para denotar a ordem de indiferença "**é indiferente a**"; e o símbolo $\succcurlyeq$ para representar ordem de preferência-indiferença "**é pelo menos tão preferível quanto**".
- Embora esses diferentes tipos de ordens de preferência ocorram naturalmente em problemas de tomada de decisão, podemos definir de forma mais compacta uma das ordens de preferência como **primitivas** e as demais como **derivadas**.
- Por exemplo, podemos começar usando a ordem preferência-indiferença $\succcurlyeq$ como **primitiva** e então definir indiferença $\sim$ e preferência estrita $\succ$ da seguinte forma:
  1. A é **indiferente** a B se e somente se A for pelo menos tão preferível quanto B e B for pelo menos tão preferível quanto A:
  $$A \sim B \quad \Leftrightarrow \quad A \succcurlyeq B \text{ e } B \succcurlyeq A$$
  2. A é **preferível** a B se e somente se A for pelo menos tão preferível quanto B e B não for pelo menos tão preferível quanto A:
  $$A \succ B \Rightarrow \quad A \succcurlyeq B \text{ e não } B \succcurlyeq A$$

- Em um problema de tomada de decisão, seja X o conjunto de todas as alternativas viáveis e sejam a, b e c quaisquer alternativas em X.

- **Em resumo:** se ρ é uma relação binária, dizemos que

  P1 ρ é **transitivo** se $a \ \rho \ b$ e $b \ \rho \ c$ implicarem $a \ \rho \ c$;

  P2 ρ é **reflexivo** se $a \ \rho \ a$;

  P3 ρ é **irreflexivo** se não valer $a \ \rho \ a$;

  P4 ρ é **simétrico** se $a \ \rho \ b$ implicar $b \ \rho \ a$;

  P5 ρ é **assimétrico** se $a \ \rho \ b$ não implicar $b \ \rho \ a$;

  P6 ρ é **antissimétrico** se $a \ \rho \ b$ e $b \ \rho \ a$ implicar $a = b$ (ou seja, a e b são idênticos).

- Para que uma relação binária se qualifique como uma relação de ordem, ela deve ser **pelo menos transitiva**.

- A maioria dos algoritmos de otimização multi-objetivo usa o conceito de **dominância**, onde duas soluções são comparadas com base em se uma domina a outra solução ou não.

- O conceito de dominância estabelece uma forma de **comparar soluções** no contexto de funções multi-objetivo, e a maioria dos métodos de otimização busca soluções não dominadas.

Dominância

Uma solução viável $\mathbf{x}^1 \in S$ é *dominante* sobre outra solução $\mathbf{x}^2 \in S$ (ou $\mathbf{x}^2$ é *dominado* por $\mathbf{x}^1$), representado por $\mathbf{x}^1 \succ \mathbf{x}^2$, se ambas as condições abaixo forem atendidas:

1. A solução $\mathbf{x}^1$ não é pior que $\mathbf{x}^2$ em todos os objetivos, ou seja, $f_k\left(\mathbf{x}^1\right) \leq f_k\left(\mathbf{x}^2\right)$, para todo $k = 1, \ldots, m_1$ e $f_k\left(\mathbf{x}^1\right) \geq f_k\left(\mathbf{x}^2\right)$, para todo $k = m_1 + 1, \ldots, m$;
2. A solução $\mathbf{x}^1$ é estritamente melhor que $\mathbf{x}^2$ em pelo menos um objetivo, ou seja, $f_\ell\left(\mathbf{x}^1\right) < f_\ell\left(\mathbf{x}^2\right)$ para algum $\ell \in \{1, \ldots, m_1\}$ ou $f_\ell\left(\mathbf{x}^1\right) > f_\ell\left(\mathbf{x}^2\right)$ por algum $\ell \in \{m_1 + 1, \ldots, m\}$.

- Se alguma dessas condições for violada, diz-se que a solução $\mathbf{x}^1$ **não domina** a solução $\mathbf{x}^2$.

- Em geral, existem **três possibilidades**:
  1. $\mathbf{x}^1$ domina $\mathbf{x}^2$;
  2. $\mathbf{x}^1$ é dominado por $\mathbf{x}^2$;
  3. não há relação de dominância entre $\mathbf{x}^1$ e $\mathbf{x}^2$.

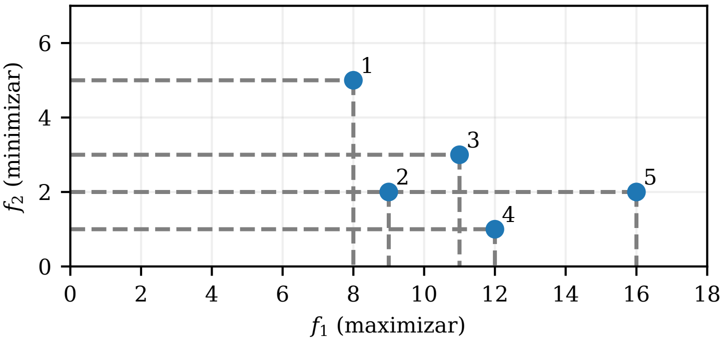
- **Exemplo:** consideremos um problema de otimização de **dois objetivos** com cinco soluções diferentes mostradas no espaço objetivo, conforme ilustrado a seguir.

- Suponhamos também que a função-objetivo 1 precisa ser **maximizada** enquanto a função objetivo 2 precisa ser **minimizada**.

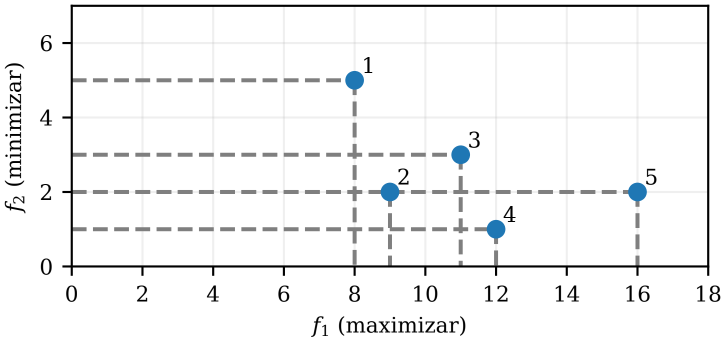
- **Cinco soluções** com diferentes valores de função objetivo são mostradas (numeradas de 1 a 5).

- Como ambas as funções objetivo são importantes para nós, geralmente é difícil encontrar uma solução que seja melhor com relação a **ambos os objetivos**.

- No entanto, podemos usar a **definição de dominância** para decidir qual solução é melhor *entre quaisquer duas soluções* dadas em termos de ambos os objetivos.



- Se as soluções 1 e 2 forem comparadas, observamos que a solução 2 é **melhor** que a solução 1 no objetivo 1 e a solução 2 também é **melhor** que a solução 1 no objetivo 2 .
- Assim, **ambas as condições** de dominação acima são satisfeitas e podemos escrever que a solução 2 **domina** a solução 1.



- Tomamos outro exemplo de comparação das soluções 2 e 5. Aqui, a solução 5 é **melhor** que a solução 2 no primeiro objetivo e a solução 5 **não é pior** que a solução 2 no segundo objetivo.
- Assim, **ambas as condições** acima para dominância também são satisfeitas e podemos escrever que a solução 5 **domina** a solução 2.

- Vamos agora discutir as diferentes **propriedades** da relação binária do operador de dominância.

Reflexiva A relação de dominância **não é reflexiva**, pois qualquer solução p não domina a si mesma de acordo com a definição. A segunda condição da relação de dominância na definição não permite que esta propriedade seja satisfeita.

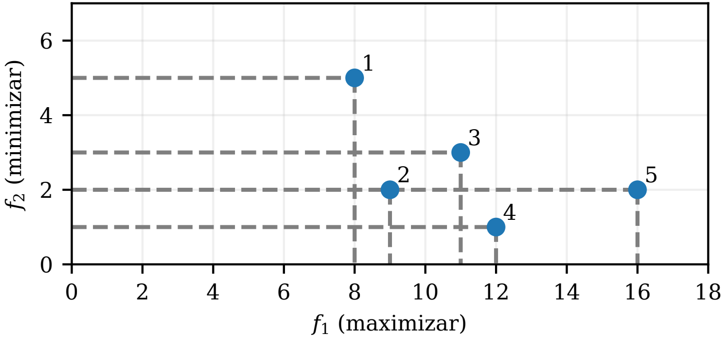
Simétrica A relação de dominância também **não é simétrica**, pois $p \preccurlyeq q$ não implica $q \preccurlyeq p$. Na verdade, o oposto é verdadeiro, ou seja, se p domina q, então q não domina p. Assim, a relação de dominância **é assimétrica**.

Antissimétrica Como a relação de dominância não é simétrica, ela também **não pode ser antissimétrica**.

Transitiva A relação de dominância **é transitiva**. Isso porque se $p \preccurlyeq q$ e $q \preccurlyeq r$, então $p \preccurlyeq r$.

- Para que uma relação binária se qualifique como uma **relação de ordem**, ela deve ser **pelo menos transitiva**.

- Como a relação de dominância não é reflexiva, é uma ordem parcial estrita (transitiva e irreflexiva).



- Continuando com as comparações do exemplo anterior, comparemos as **soluções 4 e 5**, pois essa comparação revela um aspecto interessante.
- Observamos que a solução 5 é **melhor** que a solução 4 no primeiro objetivo, enquanto a solução 5 é **pior** que a solução 4 no segundo objetivo.

Instituto Politécnico

- Assim, a primeira condição de dominância **não é satisfeita** para ambas as soluções.
- Isso simplesmente sugere que **não podemos concluir** que a solução 5 domina a solução 4, nem podemos dizer que a solução 4 domina a solução 5.
- Quando isso acontece, costuma-se dizer que as soluções 3 e 5 **são não dominadas** uma em relação à outra (lembre-se que existem três possibilidades no critério de dominância).
- Quando ambos os objetivos são importantes, **não se pode afirmar** qual das duas soluções 3 e 5 é melhor.
- **Nestes casos, como sabemos qual é a solução a ser escolhida?**
- Para um determinado conjunto finito de soluções, podemos **realizar todas as possíveis comparações de pares** e descobrir qual solução domina qual e quais soluções são não dominadas uma em relação à outra.

- No final, esperamos ter um **conjunto de soluções**, de modo que duas soluções não dominem uma à outra.
- Este conjunto também tem outra propriedade: para qualquer solução **fora desse conjunto**, sempre podemos encontrar uma solução **nesse conjunto** que dominará a primeira.
- Assim, este conjunto particular tem a propriedade de **dominar todas as outras soluções** que não pertencem a este conjunto.
- Em termos simples, isso significa que as soluções deste conjunto **são melhores** em comparação com o restante das soluções.
- Este conjunto recebe um nome especial, é chamado de **conjunto não dominado** para o conjunto de soluções dado.
- No problema anterior, quais soluções constituem o conjunto não dominado?

Atividade prática

- Vamos praticar como avaliar o conceito de **dominância**.
- Considere a definição do **problema de otimização multi-objetivo** apresentada, com $m = 2$.
- Neste problema, $f_1(x) = x^2$ e $f_2(x) = (x-1)^2$ devem ser **minimizados**.
- Gere aleatoriamente **200 pontos** com distribuição uniforme na região viável $\mathbf{x} \in [-2,\ 2]$.
- Avalie a **dominância** de cada ponto (no espaço objetivo viável).
- Plote cada ponto nos espaço das variáveis de decisão e no espaço objetivo, colorindo de vermelho os pontos **dominantes** e de azul os pontos **dominados**.

- Vamos agora nos **aprofundar** na definição do problema de otimização multi-objetivo.
- Seja $P \subset \mathbb{R}^n$ um hiperretângulo de todos os $\mathbf{x} = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$, tal que $\mathbf{x}^{\inf} \leq \mathbf{x} \leq \mathbf{x}^{\sup}$.
- Suponha que $\mathbf{f}$ é uma função definida em P com valores em $\mathbb{R}^m$, ou seja, $\mathbf{f} : P \subset \mathbb{R}^n \to \mathbb{R}^m$, com $m \geq 2$, onde $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))^\top$ é a **função multi-objetivo** e $f_r : P \to \mathbb{R}$, para $r = 1, \ldots, m$, são os **objetivos**.
- O domínio P de $\mathbf{f}$ também é chamado de **espaço das variáveis de decisão**.
- Por sua vez, o vetor $\mathbf{x} \in P$ é chamado de **vetor de decisão** e suas entradas são chamadas de **variáveis de decisão**.
- Mais especificamente, problemas típicos têm **restrições de desigualdade e igualdade**, que são escritas em forma vetorial, respectivamente, como
  - $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$, onde $\mathbf{g} : P \to \mathbb{R}^p$;
  - $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, onde $\mathbf{h} : P \to \mathbb{R}^q$.

- O conjunto de soluções que satisfazem tais restrições, denotado por

$$S = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^{\text{inf}} \leq \mathbf{x} \leq \mathbf{x}^{\text{sup}},\ \mathbf{g}\left(\mathbf{x}\right) \leq \mathbf{0},\ \mathbf{h}\left(\mathbf{x}\right) = \mathbf{0} \right\},$$

  é chamado **conjunto viável** ou **espaço de busca**, e seus pontos são denominados **soluções viáveis**.
- Um **otimizador de Pareto** não é dominado por nenhum outro ponto viável no espaço de busca.
- Mais formalmente, uma solução viável $\mathbf{x}^* \in S$ é um **otimizador de Pareto** se não houver outro vetor viável $\mathbf{x} \in S$ que domina $\mathbf{x}^*$, ou seja, $\mathbf{x} \succ \mathbf{x}^*$.
- Isso equivale a dizer que $\mathbf{x}^*$ é um elemento máximo da ordem parcial estrita $\succ$.
- Vamos denotar por $\boldsymbol{\Theta}$ o **conjunto de otimizadores de Pareto**.

- Outro conjunto relevante é a imagem da função multi-objetivo ao considerar o conjunto viável $\vartheta = \operatorname{Im}(f|_S)$, definido como

$$\operatorname{Im}(f|_S) = \{\mathbf{y} = \mathbf{f}(\mathbf{x}) \subset \mathbb{R}^m,\ \forall \mathbf{x} \in S\}$$

  e chamado **espaço objetivo**.

- Um vetor objetivo $\mathbf{z}^* \in \vartheta$ (isto é, existe $\mathbf{x}^* \in S$ em $\mathbf{z}^* = \mathbf{f}(\mathbf{x}^*)$) é um **ótimo de Pareto** se não houver outro vetor objetivo $\mathbf{z} \in \vartheta$, tal que
  - $z_k \leq z_k^*$ para cada $k = 1, \ldots, m_1$ e $z_k \geq z_k^*$ para cada $k = m_1 + 1, \ldots, m$;
  - $z_\ell < z_\ell^*$ para algum índice $\ell \in \{1, \ldots, m_1\}$ ou $z_\ell > z_\ell^*$ para algum índice $\ell \in \{m_1 + 1, \ldots, m\}$.

- Vamos denotar por $\mathcal{P}$ o **conjunto ótimo de Pareto**, também referido como **frente de Pareto** ou **curva de Pareto** (no caso em que $m = 2$).

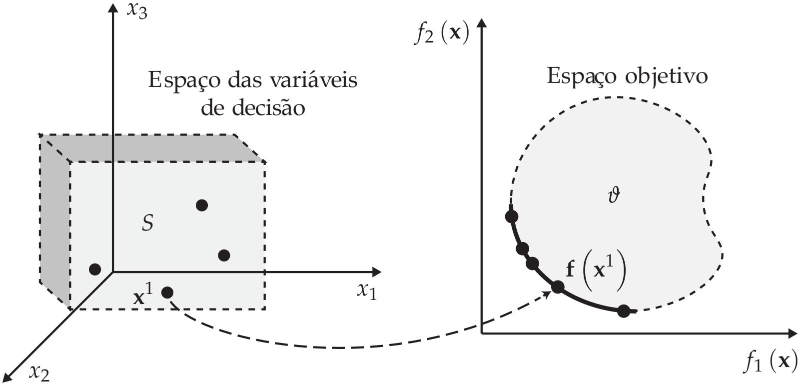
- Observe que $\mathbf{z}^*$ é um ótimo de Pareto se os vetores viáveis correspondentes, $\mathbf{x} \in S$ tal que $\mathbf{f}(\mathbf{x}) = \mathbf{z}^*$, são otimizadores de Pareto.

- Normalmente, $\mathcal{P}$ tem um **número infinito** de valores ótimos de Pareto.
- Além disso, $\boldsymbol{\Theta} = \mathbf{f}^{-1}\left(\mathcal{P}\right) = \{\mathbf{x} \in \Omega \mid \mathbf{f}\left(\mathbf{x}\right) \in \mathcal{P}\}$.
- Conhecendo todas estas definições, o problema de otimização multi-objetivo pode ser mais amplamente definido como

$$\begin{aligned} &\text{opt } \mathbf{f}\left(\mathbf{x}\right) \\ \text{Sujeito a } &\mathbf{g}\left(\mathbf{x}\right) \leq \mathbf{0} \\ &\mathbf{h}\left(\mathbf{x}\right) = \mathbf{0} \\ &\mathbf{x}^{\text{inf}} \leq \mathbf{x} \leq \mathbf{x}^{\text{sup}} \ , \end{aligned}$$

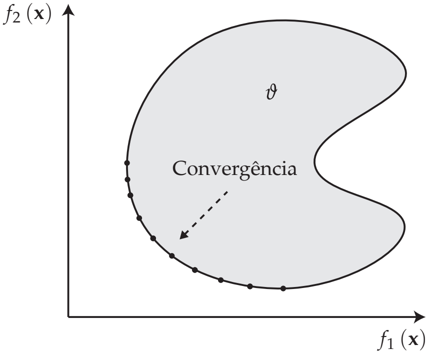
tal que sua solução é $\mathcal{P} \subset \vartheta$, o **conjunto ótimo de Pareto** no espaço objetivo e o conjunto correspondente $\boldsymbol{\Theta}$ dos **otimizadores de Pareto**.

- Exemplo de relação entre o **espaço das variáveis de decisão** e o **espaço objetivo** em um problema de otimização, onde a função multi-objetivo $\mathbf{f} : P \subset \mathbb{R}^3 \rightarrow \mathbb{R}^2$.

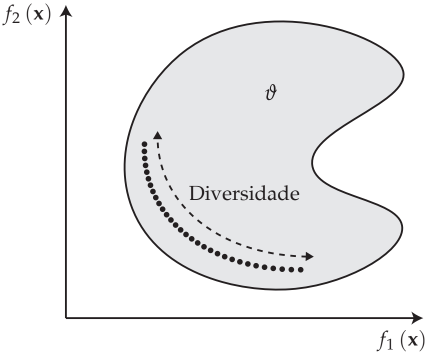


- Como $\mathcal{P}$ tem um **número infinito** de valores ótimos de Pareto, é essencial encontrar tantas soluções ótimas de Pareto quanto possível.
- Essa necessidade engloba duas exigências fundamentais no processo de otimização multi-objetivo:
  - encontrar um conjunto de valores ótimos o **mais próximo possível** da fronteira de Pareto;
  - encontrar um conjunto de valores ótimos tão **diversificadas** quanto possível.
- O primeiro requisito é obrigatório em qualquer tarefa de otimização. Por outro lado, a exigência de diversidade de resultados é totalmente **específica para otimização multi-objetivo**.
- Somente com um conjunto diversificado de soluções, pode-se ter a certeza da qualidade dos valores ótimos computados, possibilitando a **escolha adequada** de uma solução na etapa de **pós-processamento dos resultados**.

- Exemplo da relação de convergência e diversidade de soluções em um problema de $\min f_1(\boldsymbol{x})$ e $\min f_2(\boldsymbol{x})$.



Conjunto de resultados sobre a frente de Pareto, mas com diversidade insatisfatória.



Conjunto de resultados com diversidade adequada, mas afastados da frente de Pareto.