Otimização não-linear

Prof. Tarcisio F. Maciel, Dr.-Eng.
Colaboradores: Diego A. Sousa, M.-Eng., José Mairton B. da Silva Jr., M.-Eng., Francisco Hugo C. Neto, M.-Eng., e Yuri Victor L. de Melo, M.-Eng.

Universidade Federal do Ceará
Centro de Tecnologia
Programa de Pós-Graduação em Engenharia de Teleinformática

04/07/2018

Introdução

Conteúdo

Conteúdo

Introdução [NW06]

- Pessoas otimizam
  - Investidores ➜ criam portfólios que minimizam riscos e atingem uma certa taxa de retorno
  - Fabricantes ➜ maximizam eficiência no projeto e operação de suas plantas produtivas
  - Engenheiros ➜ ajustam parâmetros reduzir custos e aumentar a eficiência de seus projetos
- A natureza otimiza
  - Sistemas físicos ➜ tendem ao estado de mínima energia
  - Raios de luz ➜ percorrem o caminho de menor tempo de percurso
  - Moléculas ➜ se acomodam para minimizar a energia potencial dos elétrons
- Otimização é uma ferramenta importante para tomada de decisões e análise de sistemas físicos

Introdução [NW06]

- O que é otimização?
  - Dar a algo um rendimento ótimo, criando-lhe as condições mais favoráveis ou tirando o melhor partido possível; tornar algo ótimo ou ideal [dHF10]
- Porque otimizar?
  - Com otimização é possível melhorar o desempenho de um sistema, ou seja, deixar o sistema mais rápido e eficiente [NW06]
- Como otimizar?
  - Para otimizar é preciso definir o objetivo, uma medida que quantifica o desempenho do sistema
  - O objetivo depende de certos de certas características do sistema, chamadas de variáveis que otimizam o sistema
  - Por fim é frequente o uso de restrições que descrevem situações do sistema consideradas não desejáveis [NW06]

Conteúdo

Elementos de um problema de otimização

- Uso da ferramenta de otimização ➜ identificar alguns elementos
  - Objetivo ➜ medida quantitativa do desempenho do sistema em estudo ➜ lucro, tempo, energia, entre outros, ou uma combinação de fatores resultando em um escalar
  - Variáveis ➜ parâmetros cujos valores podem ser ajustados e dos quais depende o desempenho do sistema
  - Restrições ➜ condições às quais os valores da variáveis estão sujeitos
- Modelagem de problemas de otimização ➜ Processo de identificação do objetivo, variáveis e restrições
  - Construção de um modelo apropriado ➜ algumas vezes é o passo mais importante
  - Excessivamente simplista ➜ não provê informação suficiente
  - Complexo demais ➜ difícil de resolver
- Após modelado ➜ solucionado usando um algoritmo de otimização
- Não há algoritmo universal ➜ coleção de algoritmos especializados para cada tipo de problema
- Maiores benefícios ➜ surgem quando o tipo do problema é conhecido
- Resultado do algoritmo ➜ validado utilizando as condições de optimalidade

Modelo matématico de um problema de otimização

- Matematicamente, otimização é a maximização ou minimização de uma função objetivo sujeita a restrições sobre suas variáveis de otimização
- Em nossos modelos matemáticos, tipicamente:
  - $\boldsymbol{x} \in \mathbb{R}^n$ ➜ vetor de variáveis de otimização
  - $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ ➜ função objetivo que se deseja maximizar ou minimizar
  - $f_i(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ ➜ as restrições de igualdade e desigualdade que $\boldsymbol{x}$ deve satisfazer
- Logo, pode-se escrever um problema de otimização como

$$\boldsymbol{x}^\star = \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} \; f(\boldsymbol{x}) \tag{1a}$$

$$\text{sujeito a} \quad f_i(\boldsymbol{x}) = 0, \quad i = 1, 2, \ldots, I, \tag{1b}$$

$$f_j(\boldsymbol{x}) \geq 0, \quad j = 1, 2, \ldots, J, \tag{1c}$$

onde i e j são os índices para as restrições de igualdade e desigualdade, respectivamente, e $\boldsymbol{x}^\star$ é uma solução ótima

Exemplo de problema de otimização: transporte de produtos [NW06]

- Uma companhia possui 02 fábricas F_1 e F_2 e 12 lojas $R_1, R_2, \ldots, R_{12}$. Cada fábrica F_i pode produzir a_i toneladas de um produto (a_i é a capacidade da planta de produção) e cada loja possui uma demanda semanal de b_j toneladas do produto. O custo de transporte da fábrica F_i para a loja R_j de uma tonelada do produto é $c_{i,j}$. O problema é determinar as quantidades $x_{i,j} \in \mathbb{R}_+$ do produto que devem ser transportadas de cada fábrica para cada loja de modo a atender a todos os requisitos e minimizar o custo total.

Exemplo de problema de otimização: transporte de produtos [NW06]

- Uma companhia possui 02 fábricas F_1 e F_2 e 12 lojas $R_1, R_2, \ldots, R_{12}$. Cada fábrica F_i pode produzir a_i toneladas de um produto (a_i é a capacidade da planta de produção) e cada loja possui uma demanda semanal de b_j toneladas do produto. O custo de transporte da fábrica F_i para a loja R_j de uma tonelada do produto é $c_{i,j}$. O problema é determinar as quantidades $x_{i,j} \in \mathbb{R}_+$ do produto que devem ser transportadas de cada fábrica para cada loja de modo a atender a todos os requisitos e minimizar o custo total. Esse problema pode ser formulado como segue:

$$\{x^\star_{i,j}\} = \underset{\{x_{i,j}\}}{\text{minimize}} \sum_{i=1}^{2} \sum_{j=1}^{12} c_{i,j} x_{i,j} \tag{2a}$$

$$\text{sujeito a} \quad \sum_{j=1}^{12} x_{i,j} \leq a_i, \quad i = 1, 2 \tag{2b}$$

$$\sum_{i=1}^{2} x_{i,j} \geq b_j, \quad j = 1, 2, \ldots, 12 \tag{2c}$$

$$x_{i,j} \geq 0, \quad i = 1, 2 \quad \text{e} \quad j = 1, 2, \ldots, 12 \tag{2d}$$

- Este problema é um problema de otimização linear ➜ função custo e todas as restrições são funções lineares das variáveis do problema

Reescreva o problema acima em forma vetorial/matricial

Exemplo de problema de otimização: minimização da soma das correlações espaciais [MK06]

- A correlação espacial $\rho_{i,j}$ entre os canais $\boldsymbol{h}_i = \begin{bmatrix} h_{i,1} & h_{i,2} & \dots & h_{i,N} \end{bmatrix}$ e $\boldsymbol{h}_j = \begin{bmatrix} h_{j,1} & h_{j,2} & \dots & h_{j,N} \end{bmatrix}$ do enlace direto de uma estação rádio base com N antenas para os terminais móveis i, j é dada por $\rho_{i,j} = \frac{|\boldsymbol{h}_i \boldsymbol{h}_j^{\mathrm{H}}|}{\|\boldsymbol{h}_i\|_2 \|\boldsymbol{h}_j\|_2}$. Sabendo que existem K terminais móveis, selecione $G \leq N$ terminais móveis tal que a soma das correlações entre eles dois-a-dois seja mínima, ou seja, selecione os G terminais móveis com os canais menos correlacionados.

Exemplo de problema de otimização: minimização da soma das correlações espaciais [MK06]

- A correlação espacial $\rho_{i,j}$ entre os canais $\boldsymbol{h}_i = \begin{bmatrix} h_{i,1} & h_{i,2} & \ldots & h_{i,N} \end{bmatrix}$ e $\boldsymbol{h}_j = \begin{bmatrix} h_{j,1} & h_{j,2} & \ldots & h_{j,N} \end{bmatrix}$ do enlace direto de uma estação rádio base com N antenas para os terminais móveis i, j é dada por $\rho_{i,j} = \dfrac{|\boldsymbol{h}_i \boldsymbol{h}_j^{\mathrm{H}}|}{\|\boldsymbol{h}_i\|_2 \|\boldsymbol{h}_j\|_2}$. Sabendo que existem K terminais móveis, selecione $G \leq N$ terminais móveis tal que a soma das correlações entre eles dois-a-dois seja mínima, ou seja, selecione os G terminais móveis com os canais menos correlacionados. Esse problema pode ser formulado como segue:

$$\boldsymbol{x}^{\star} = \underset{\boldsymbol{x}}{\text{minimize}} \ \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{x}, \tag{3a}$$

$$\text{sujeito a} \quad \mathbf{1}^{\mathrm{T}}\boldsymbol{x} = G, \tag{3b}$$

$$\boldsymbol{x} \in \mathbb{B}^{K}, \tag{3c}$$

onde $\boldsymbol{R} = [\rho_{i,j}]_{i,j}, \quad i, j \in \{1, 2, \ldots, K\}$.

- Este problema é um problema de otimização binário quadrático ➜ função custo quadrática com variáveis de otimização binárias

Reescreva o problema acima utilizando somatórios

Conteúdo

Classes de problemas de otimização

- Natureza das variáveis de otimização, da função objetivo, e das restrições ➜ diferentes tipos de problemas de otimização e algoritmos de otimização
- Variáveis de otimização
  - $\boldsymbol{x} \in \mathbb{R}^n$ ➜ otimização contínua (mais fácil de resolver)
  - $\boldsymbol{x} \in \mathbb{Z}^n$ ➜ otimização inteira (pode requerer relaxações contínuas)
  - $x_1, \ldots, x_k \in \mathbb{R}$ e $x_{k+1}, \ldots, x_n \in \mathbb{Z}$ ➜ otimização inteira mista
- Função objetivo e restrições
  - $f(\boldsymbol{x})$, $f_i(\boldsymbol{x})$ e $f_j(\boldsymbol{x})$ lineares ➜ Otimização linear
  - $f(\boldsymbol{x})$ e $f_i(\boldsymbol{x})$ convexas, $f_j(\boldsymbol{x})$ lineares ➜ Otimização convexa
- Máximo/mínimo local ➜ $f(\boldsymbol{x}^\star)$ é um máximo/mínimo local de $f(\boldsymbol{x})$ se existe um sub-espaço aberto $\mathbb{A} \subset \mathbb{R}^n$ contendo $\boldsymbol{x}^\star$ tal que $f(\boldsymbol{x}) \lesseqqgtr f(\boldsymbol{x}^\star)$, $\forall \boldsymbol{x} \in \mathbb{A}$.
- Máximo/mínimo global ➜ $f(\boldsymbol{x}^\star)$ é um máximo/mínimo global de $f(\boldsymbol{x})$ se $f(\boldsymbol{x}) \lessgtr f(\boldsymbol{x}^\star)$, $\forall \boldsymbol{x} \in \mathbb{R}^n$

Classes de problemas de otimização

- Algoritmos de otimização ➜ especializados para cada tipo de problema
  - Otimização linear ➜ método Simplex
  - Otimização convexa ➜ método dos pontos interiores
  - Otimização linear inteira ➜ método *branch-and-bound*
  - Otimização sem restrições ➜ método de busca direta e métodos do gradiente
  - Otimização com restrições ➜ métodos dos pontos interiores
  - Otimização determinística ➜ restrições e parâmetros dados por funções bem definidas
  - Otimização estocástica ➜ restrições ou parâmetros dependem de variáveis aleatórias

Exemplo de problema de otimização: mínimos quadrados [BV04]

- Considere o problema de minimizar a soma dos erros quadráticos entre as componentes de um vetor $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ e um vetor de referência $\boldsymbol{b}$. Esse é um problema de mínimos quadrados sem restrições que pode ser escrito como

$$\boldsymbol{x}^\star = \underset{\boldsymbol{x}}{\text{minimize}}\ \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \tag{4}$$

Exemplo de problema de otimização: mínimos quadrados [BV04]

- Considere o problema de minimizar a soma dos erros quadráticos entre as componentes de um vetor $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ e um vetor de referência $\boldsymbol{b}$. Esse é um problema de mínimos quadrados sem restrições que pode ser escrito como

$$\boldsymbol{x}^\star = \underset{\boldsymbol{x}}{\text{minimize}} \ \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \tag{4}$$

- Note que

$$\begin{aligned} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 &= (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{b}^{\mathrm{T}})(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \\ &= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} \end{aligned} \tag{5}$$

- Derivando a equação acima em relação a $\boldsymbol{x}$ e igualando a $\boldsymbol{0}$ temos

$$\frac{d}{d\boldsymbol{x}}\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}\right) = 0 \Rightarrow 2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} = 0 \Rightarrow \boxed{\boldsymbol{x}^\star = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}} \tag{6}$$

Liste as classes de problemas de otimização às quais esse problema pertence

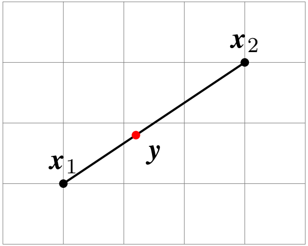
Conjuntos convexos e funções convexas

Conteúdo

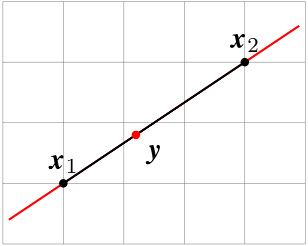
Conteúdo

Conjuntos afins

- Linha e segmentos
  - Sejam os pontos $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ com $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ ➜ pontos $\boldsymbol{y}$ da forma $\boldsymbol{y} = \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2$ com $\theta \in \mathbb{R}, 0 \leq \theta \leq 1$, formam um segmento fechado ligando $\boldsymbol{x}_1$ a $\boldsymbol{x}_2$
  - Representando $\boldsymbol{y} = \boldsymbol{x}_2 + \theta(\boldsymbol{x}_1 - \boldsymbol{x}_2)$ ➜ Ponto base $\boldsymbol{x}_2$ e direção $(\boldsymbol{x}_1 - \boldsymbol{x}_2)$ apontando de $\boldsymbol{x}_2$ para $\boldsymbol{x}_1$ e escalonada por θ

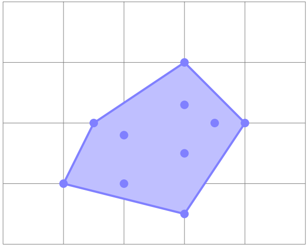


- Conjuntos afins
  - Um conjunto $\mathbb{A} \subset \mathbb{R}^n$ é afim ➜ para qualquer $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{A} \subset \mathbb{R}$ e $\theta \in \mathbb{R}$, o segmento de reta $\boldsymbol{y} = \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in \mathbb{A}$
  - Combinação afim ➜ $\boldsymbol{y} = \theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2 + \ldots + \theta_k \boldsymbol{x}_k \in \mathbb{A} \subset \mathbb{R}$, com $\theta_1 + \theta_2 + \ldots + \theta_k = 1$

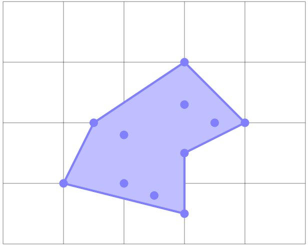


Conjuntos convexos

- Conjuntos convexos
  - Um conjunto $\mathbb{A} \subset \mathbb{R}^n$ é convexo caso um segmento entre $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{A}$ e $0 \leq \theta \leq 1$, pertença a $\mathbb{A}$, ou seja, $\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in \mathbb{A}$
  - Em termos geométricos, um conjunto convexo é um conjunto sem buracos ou reentrâncias



- Invólucro convexo (*convex hull*)
  - Dado um conjunto $\mathbb{A}$, o invólucro convexo de $\mathbb{A}$ é o menor conjunto convexo que engloba $\mathbb{A}$, sendo denotado por $\operatorname{conv}\{\mathbb{A}\} = \{\theta_1 \boldsymbol{x}_1 + \cdots + \theta_k \boldsymbol{x}_k \mid \boldsymbol{x}_i \in \mathbb{A}, \theta_i \geq 0, i = 1, \cdots, k, \theta_1 + \cdots + \theta_k = 1\}$



Conteúdo

Operações sobre conjuntos que preservam convexidade

- Interseção ➜ se $\mathbb{A}_\alpha$ é convexo para todo $\alpha \in \mathbb{R}$, então

$$\bigcap_{\alpha \in \mathbb{R}} \mathbb{A}_\alpha \qquad \text{também é convexo} \tag{7}$$

- Multiplicação por escalar ➜ se $\mathbb{A} \subseteq \mathbb{R}^n$ é convexo e $\alpha \in \mathbb{R}$, então

$$\alpha\mathbb{A} = \{\alpha\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{A}\} \qquad \text{também é convexo} \tag{8}$$

- Translação ➜ se $\mathbb{A} \subseteq \mathbb{R}^n$ é convexo e $\boldsymbol{\alpha} \in \mathbb{R}^n$, então

$$\mathbb{A} + \boldsymbol{\alpha} = \{\boldsymbol{x} + \boldsymbol{\alpha} \mid \boldsymbol{x} \in \mathbb{A}\} \qquad \text{também é convexo} \tag{9}$$

- Soma ➜ se $\mathbb{A}_1$ e $\mathbb{A}_2$ são convexos, então

$$\mathbb{A}_1 + \mathbb{A}_2 = \{\boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} \in \mathbb{A}_1, \boldsymbol{y} \in \mathbb{A}_2\} \qquad \text{também é convexo} \tag{10}$$

- Produto cartesiano ➜ se $\mathbb{A}_1$ e $\mathbb{A}_2$ são convexos, então

$$\mathbb{A}_1 \times \mathbb{A}_2 = \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{x} \in \mathbb{A}_1, \boldsymbol{y} \in \mathbb{A}_2\} \qquad \text{também é convexo} \tag{11}$$

Conteúdo

Hiperplanos, cones e desigualdades generalizadas

- Hiperplano ➜ conjunto de pontos, podendo ser escrito como

$$\{\boldsymbol{x} \mid \boldsymbol{a}^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}_0) = 0)\} \tag{12}$$

onde $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{a} \neq \boldsymbol{0}$, e $\boldsymbol{x}_0$ determina o offset do hiperplano. Um hiperplano divide o espaço em dois semi-espaços

- Semi-espaço ➜ conjunto da forma

$$\{\boldsymbol{x} \mid \boldsymbol{a}^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}_0) \leq 0)\} \tag{13}$$

onde $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{a} \neq \boldsymbol{0}$

TODO

Adicionar figuras

Hiperplanos, cones e desigualdades generalizadas

- Cone é conjunto de pontos tais que $\forall \boldsymbol{x} \in \mathbb{A} \subset \mathbb{R}^n$ e $\theta \geq 0, \theta \in \mathbb{R}_+$

$$\theta \boldsymbol{x} \in \mathbb{A}, \forall \boldsymbol{x} \in \mathbb{A} \tag{14}$$

- Cone Convexo é um conjunto que é simultaneamente um cone e convexo, ou seja, para qualquer $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{A} \subset \mathbb{R}^n$ e $\theta_1, \theta_2 \geq 0, \theta_1$

$$\theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2 \in \mathbb{A} \tag{15}$$

TODO

Adicionar figuras

Conjuntos Convexos

- A bola Euclidiana com o centro em $\boldsymbol{x}_c$ e raio r é representada por

$$B(\boldsymbol{x}_c, r) = \{\boldsymbol{x}_c + r\boldsymbol{u} \mid \|\boldsymbol{u}\|_2 \leq 1\} \tag{16}$$

- O elipsoide com o centro em $\boldsymbol{x}_c$ é um conjunto representado na forma

$$\mathcal{E} = \{\boldsymbol{x} \mid (\boldsymbol{x} - \boldsymbol{x}_c)^{\mathrm{T}} \boldsymbol{P}^{-1} (\boldsymbol{x} - \boldsymbol{x}_c) \leq 1)\} \tag{17}$$

onde $\boldsymbol{P} \in \mathbb{A} \subset \mathbb{S}^n_{++}$, ou seja, $\boldsymbol{P}$ é uma matriz simétrica positiva definida

TODO

Adicionar figuras

Conjuntos Convexos

- Poliedro
  - O poliedro é definido como um conjunto de igualdades lineares e inequações.

$$Ax \preceq b, Cx = d \tag{18}$$

    onde $\preceq$ é o símbolo que representa desiguadade componente a componente entre vetores.
  - Dessa forma, a representação do poliedro pode escrita como:

$$P = \{x \mid Ax \preceq b, Cx = d\} \tag{19}$$

Conjuntos Convexos

- Norma da Bola com centro x_c e raio r é representado por:

$$\{x \mid \|x - x_c\| \leq r\} \tag{20}$$

onde $\|\cdot\|$ é a norma.

- A norma do cone pode ser definido como:

$$\{(x, t) \in \mathbb{R}^n + 1 \mid \|x\| \leq t\} \tag{21}$$

Conjuntos Convexos

- PSD (Cone positivo semidefinido) é um cone formado a parti das matrizes positivas semidefinidas que utiliza a notação:

$$\mathbb{A}_+^n = \{X \in \mathbb{A}^n \mid X \succ 0\} \tag{22}$$

onde $\mathbb{A}^n$ denota o conjunto de matrizes simetricas. Dessa forma o conjunto $\mathbb{A}_+^n$ é um cone positivo semidefinido, se $\theta_1, \theta_2 \geqslant 0$ e $A, B \in \mathbb{A}_+^n$, assim $\theta_1 A + \theta_2 B \in \mathbb{A}_+^n$.

Conjuntos Convexos

- Desigualdade generalizadas
  - Um cone pode ser usado para definir a desigualdade generalizada, que é similar a relação de ordem apresentada em $\mathbb{R}$. Desta forma, para um cone K onde $K \subseteq \mathbb{R}^n$, as desigualdades não estritas e estritas $\preceq_K$ e $\prec_K$, respectivamente são definidas da seguinte maneira:

$$x \preceq_K y \Leftrightarrow y - x \in K \tag{23a}$$

$$x \prec_K y \Leftrightarrow y - x \in \mathbf{int}(K) \tag{23b}$$

    onde $\mathbf{int}(K)$ representa o interior do conjunto K.
  - Propriedades
    - Se $x \preceq_K y$ e $u \preceq_K v$, então $x + u \preceq_K y + v$
    - Se $x \preceq_K y$ e $y \preceq_K u$, então $x \preceq_K u$
    - Se $x \preceq_K y$ e $y \preceq_K x$, então $x = y$
    - Se $x \prec_K y$ e $u \preceq_K v$, então $x + u \prec_K y + v$

Conteúdo

Conteúdo

Propriedades básicas

- Uma função $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ é convexa somente se o domínio de f $(\mathrm{dom}\ f)$ for convexo e $\forall x, y \in \mathrm{dom}\ f$, e $0 \leq \theta \leq 1$ (24) for verdade.

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \tag{24}$$

- Observando a condição (24), temos que em uma função $f(\cdot)$ convexa, qualquer segmento de reta entre os pontos $(x, f(x))$ e $(y, f(y))$ estará acima de $f(\cdot)$ entre os valores x e y, como mostrado na figura abaixo.

Propriedades básicas

- Dizemos que uma função $f(\cdot)$ é estritamente convexa se a desigualdade ocorrer sempre para $0 < \theta < 1$ e $x \neq y$, ou seja,

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y) \tag{25}$$

- Define-se uma função $f(\cdot)$ como côncava, se $-f(\cdot)$ for convexa.
- Funções afins são convexas e côncavas, visto que

$$f(\theta x + (1-\theta)y) = \theta f(x) + (1-\theta)f(y) \tag{26}$$

para $0 \leq \theta \leq 1$.

Propriedades básicas

- Uma função $f(\cdot)$ é convexa se e somente se ela é convexa quando restrita a qualquer linha que intercepta seu domínio. Ou seja, $f(x)$ é convexa se e somente se para todo $x \in \mathrm{dom}\, f$ e para todo v, a função $g(t) = f(x + tv)$ é convexa no domínio $\{t | x + tv \in \mathrm{dom}\, f\}$.
- Esta propriedade é muito importante, pois pode-se avaliar a convexidade de uma função, restringindo-a a uma linha.

Extensão

- As vezes é conveniente que o domínio da função $f(\cdot)$ seja estendido para todo $\mathbb{R}^n$, atribuindo o valor ∞ para $f(x)$ quando $x \notin \operatorname{dom} f$.
- Assim, a extensão de $f(x)$, $\tilde{f} : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \infty$ é definida como:

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \operatorname{dom} f \\ \infty & x \notin \operatorname{dom} f \end{cases} \tag{27}$$

- Essa extensão é útil pelo fato de simplificar a notação e de não haver a necessidade de se explicitar o seu domínio.
- De modo análogo, podemos estender funções côncavas adicionando $-\infty$ a seu domínio.

Condições de primeira ordem

- Seja $f(\cdot)$ uma função diferenciável, então $f(\cdot)$ é convexo se e somente se o $\mathrm{dom}\, f$ for convexo e

$$f(y) \geq f(x) + \nabla f(x)^{\mathrm{T}}(y - x) \quad \forall x, y \in \mathrm{dom}\, f \tag{28}$$

- Note que o segundo termo de (28) é uma função afim em relação a y e que é igual a aproximação de primeira ordem de Taylor de $f(y)$ em torno de x.

Condições de primeira ordem

- Assim, se a aproximação de primeira ordem de Taylor de uma função sempre representar um subestimador global, a função $f(\cdot)$ é convexa.
- De (28), podemos ver que a partir da informação local (x e $\nabla f(x)$) nós podemos derivar a informação global (subestimador global) de $f(\cdot)$.
- Podemos ainda deduzir de (28) que se $\nabla f(x) = 0$, então $f(y) \geq f(x)$, o que implica dizer que o ponto $(x, f(x))$ representa o mínimo global de $f(\cdot)$.
- Para que $f(\cdot)$ seja estritamente convexo, temos que a desigualdade em (28) tem que ser estrita, ou seja,

$$f(y) > f(x) + \nabla f(x)^{\mathrm{T}}(y - x) \quad \forall x, y \in \operatorname{dom} f \text{ e } x \neq y \tag{29}$$

Condições de segunda ordem

- Seja $f(\cdot)$ duas vezes diferenciável. $f(\cdot)$ é convexo se e somente se o $\operatorname{dom} f$ é convexo e sua Hessiana é semi-positiva definida $\forall x \in \operatorname{dom} f$, ou seja,

$$\nabla^2 f(x) \succeq 0. \tag{30}$$

- No caso das funções reais (30) se reduz a $f''(x) \geq 0$, o que implica dizer que $f(\cdot)$ é não decrescente.
- Geograficamente, a condição (30) pode ser interpretada como o gráfico da função possuir curvatura sempre positiva.
- De modo análogo, temos que $f(\cdot)$ é côncava se e somente se

$$\nabla^2 f(x) \preceq 0. \tag{31}$$

- Se $\nabla^2 f(x) \succ 0$, então $f(\cdot)$ é estritamente convexa, contudo, a recíproca não é verdadeira!

Conjunto de subníveis

- O conjunto de subnível α de uma função $f : \mathbb{R}^n \to \mathbb{R}$ é definido como:

$$C_\alpha = \{x \in \operatorname{dom} f | f(x) \leq \alpha\} \tag{32}$$

- Conjuntos de subnível de uma função convexa são convexos para qualquer α.
- Prova: Seja $x, y \in \mathbb{A}_\alpha$, então $f(x) \leq \alpha$ e $f(y) \leq \alpha$. Assim, de (24), temos que:

$$\begin{aligned} f(\theta x + (1-\theta)y) &\leq \theta f(x) + (1-\theta) f(y) \\ &\leq \theta\alpha + (1-\theta)\alpha \leq \alpha \end{aligned}$$

  Logo, $\theta x + (1-\theta)y \in \mathbb{A}_\alpha$
- A recíproca não é verdadeira! Uma função pode ter todos seus subníveis convexos e não ser convexa.

Epígrafo e Hipografo

- O grafo de uma função $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ é definido como

$$\{(x, f(x)) | x \in \mathrm{dom}\, f\}, \tag{33}$$

  que é um subconjunto de $\mathbb{R}^{n+1}$.

- O epígrafo de $f(\cdot)$ é definido como:

$$\mathrm{epi}\, f = \{(x, t) | x \in \mathrm{dom}\, f, f(x) \leq t\} \tag{34}$$

  que é um subconjunto de $\mathbb{R}^{n+1}$.

Epígrafo e Hipografo

- O epígrafo faz a ligação entre funções e conjuntos convexos.
- Uma função $f(\cdot)$ é convexa se e somente se seu epígrafo representar um conjunto convexo.
- O hipografo de uma função é definido como sendo

$$\text{hypo } f = \{(x,t) | x \in \text{dom } f, t \leq f(x)\} \tag{35}$$

- A função $f(\cdot)$ é côncava se e somente seu hipografo for convexo.

Desigualdade de Jensen e extensões

-

Conteúdo

Conteúdo

Conteúdo

Conteúdo

Conteúdo

Conteúdo

Bibliografia

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, 1st ed. Cambridge University Press, 2004.

[dHF10] A. B. de Holanda Ferreira, *Dicionário do Aurélio*, 5th ed. Positivo Editora, 2010.

[MK06] T. F. Maciel and A. Klein, "A low-complexity SDMA grouping strategy for the downlink of Multi-User MIMO systems," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, sep 2006.

[NW06] J. Nocedal and S. J. Wright, *Numerical optimization*, 2nd ed., T. V. Mikosh, S. M. Robinson, and S. I. Resnick, Eds. Springer, 2006.