Lista 1: Otimização I

A. Ramos *

August 15, 2017

Abstract

Lista em constante atualização.

1. Funções com valores extendidos, semicontinuidade inferior
2. Existência de soluções, coercividade
3. Convexidade

1. Considere a função $f : \mathbb{R}^n \to \mathbb{R}$ definida como $f(x) = x^T Ax + b^T x + c$, com A simetrico e $b \neq 0$.
   - Mostre que f admite solução se, e somente se A é definida positiva.
2. Mostre a existencia de autovalores para matrices simetricas. Para isso, faça o seguinte:
   (a) Seja $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ uma função contínua tal que $f(\alpha x) = f(x)$, para todo $\alpha > 0$ e $x \neq 0$. Mostre que f admite solução.
   (b) Seja A uma matriz simetrica e defina $f(x) := \frac{x^T Ax}{\|x\|^2}$, para $x \neq 0$. Prove que f tem solução
   (c) Calcule $\nabla f(x)$, mostre que as soluções são os autovetores de A. Qual é a relação dos autovalores com $\inf f(x)$?
3. Denote por $\mathrm{Sym}^m(\mathbb{R})$ o conjunto das matrizes reais simetricas de ordem $m \times m$, por $\mathrm{Sym}^m_+(\mathbb{R})$ o conjunto das reais matrizes simetricas de $m \times m$ semidefinida positiva e por $\mathrm{Sym}_{++}(\mathbb{R})$ o conjunto de matrizes reais simetricas de $m \times m$ definidas positivas.
   - Prove que $\mathrm{Sym}^m(\mathbb{R})$ é um espaço de Hilbert, considerando o seguinte produto interno $\langle X, Y \rangle := \mathrm{tr}(XY)$, onde $\mathrm{tr}(A)$ é o traço da matriz A.
   - Mostre que $\mathrm{Sym}^m_+(\mathbb{R})$ is a cone fechado convexo, cujo interior é $\mathrm{Sym}^m_{++}(\mathbb{R})$.

   Seja $C \in \mathrm{Sym}^m_{++}(\mathbb{R})$, $\eta > 0$, $A_i \in \mathrm{Sym}^m_+(\mathbb{R})$, e $b_i \in \mathbb{R}$, para todo $i = 1, \ldots, n$. Mostre que o problema de minimizar

   $$\text{minimizar } \langle X, C \rangle + \eta\beta(X) \text{ sujeito a } X \in \mathrm{Sym}^m_{++}(\mathbb{R}), \text{ e } \langle X, A_i \rangle = b_i, \forall i$$

   admite solução, onde

   $$\beta(X) = -\ln\det(X) \text{ para } X \in \mathrm{Sym}^m_{++}(\mathbb{R}) \quad (\ \beta(X) \text{ é chamado de } \textit{barreira funcional} \ )$$

   Para isso prove que os conjunto de niveis de $\langle X, C \rangle + \eta\beta(X)$ são compactos, e concluia. Um roteiro é o seguinte:
   (a) Seja $c \in \mathbb{R}_+$. Mostre que $\beta(t) := ct - \ln(t), t > 0$ tem conjuntos de niveis compacto.
   (b) Verifique que $\|X\|^2 = \mathrm{tr}(X^2) = \sum_{i=1}^m \lambda_i^2(X)$, onde $\lambda_1(X) \leq \lambda_2(X) \leq \cdots \leq \lambda_m(X)$ são os autovalores (com multiplicidade) da matriz X.
   (c) Seja $X \in \mathrm{Sym}^m(\mathbb{R})$ e denote por $\lambda(X)$ o vetor em $\mathbb{R}^m$ cujas componentes são $\lambda_i(X)$, $i = 1, \ldots, m$. Use a *desigualdade de Fan* [1], para encontrar um limitante inferior para $\langle X, C \rangle = \mathrm{tr}(XC)$ em função dos autovalores de X e C.
   (d) Use os itens anteriores para mostrar que conjunto de niveis de $\langle X, C \rangle + \eta\beta(X)$ são compactos.

   Esse tipo de problema aparece naturalmente quando estudamos o método de ponto interiores para programação semidefinida positiva (SDP programming).

*Department of Mathematics, Federal University of Paraná, PR, Brazil. Email: `albertoramos@ufpr.br`.

[1] Sempre temos que $\mathrm{tr}(XY) \leq \lambda(X)^T\lambda(Y)$ para todo $X, Y \in \mathrm{Sym}^m(\mathbb{R})$. A igualdade vale se, e somente se existe uma matriz ortogonal U tal que $UXU^T = Diag\lambda(X)$ e $UYU^T = Diag\lambda(Y)$. A desigualdade de Fan pode ser interpretada como um refinamento da desigualdade de Cauchy-Schwarz.

4. Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função derivável tal que

$$\lim_{\|x\| \to \infty} \frac{f(x)}{\|x\|} = \infty.$$

Mostre que para todo $y \in \mathbb{R}^n$, existe um $x \in \mathbb{R}^n$ tal que $\nabla f(x) = y$. Em outras palavras, $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ é surjetiva.

5. Sejam $f, g : (X, \tau) \to \mathbb{R} \cap \{+\infty\}$ funções com valores extendidos e $x^* \in X$. Mostre que

$$\liminf_{x \to x^*} (f(x) + g(x)) \geq \liminf_{x \to x^*} f(x) + \liminf_{x \to x^*} g(x), \text{ se a soma da direita não é } \infty - \infty.$$

6. Seja (X, τ) um espaço topológico.
   (a) Mostre que δ_C é τ-lsc se, e somente se C é τ-fechado.
   (b) Mostre que epi(f) é fechado na topologia produto $X \times \mathbb{R}$ se, e somente se $f : X \to \mathbb{R} \cup \{+\infty\}$ é τ-lsc.
   (c) Mostre que $\sup\{f_i : i \in I\}$ é τ-lsc se $f_i : X \to \mathbb{R} \cup \{+\infty\}$ é τ-lsc, $\forall i \in I$.
   Em particular, mostre que o supremo de funções τ-contínuas é no máximo τ-lsc, fornecendo um exemplo onde o supremo de funções contínuas não é continua.
   (d) Prove que $\sum_{i=1}^{m} \alpha_i f_i$ é τ-lsc se f_i é τ-lsc e $\alpha_i \in \mathbb{R}_+$, $\forall i$. Isto é, a combinação positiva de funções τ-lsc é τ-lsc.

7. Seja (X, d) um espaço métrico e $f : (X, d) \to \mathbb{R} \cup \{+\infty\}$. Mostre que f é lsc se para todo $(x, \lambda) \in X \times \mathbb{R}$ temos que $f(x) \leq \lambda$ sempre que $(x^k, f(x^k)) \to (x, \lambda)$.

8. *Principio Variational de Ekeland.* O seguinte teorema é uma pequena variação de teorema apresentado em aula.
   Seja (X, d) um espaço metrico completo, $f : X \to \mathbb{R} \cap \{+\infty\}$ uma função lsc e limitada inferiormente.
   Suponha que existe $\varepsilon > 0$ e $z \in X$ tal que
   $$f(z) \leq \inf f + \varepsilon.$$
   Então, para todo $\lambda > 0$, *existe* um elemento $z_\lambda \in X$ com as seguintes propriedades:
   (a)
   $$d(z_\lambda, z) \leq \lambda, \qquad f(z_\lambda) + \frac{\varepsilon}{\lambda} d(z_\lambda, z) \leq f(z)$$
   (b) Para todo $y \neq z_\lambda$, temos que
   $$f(z_\lambda) < f(y) + \frac{\varepsilon}{\lambda} d(z_\lambda, y)$$

   Para provar o teorema use os seguinte passos. A ideia é construir indutivamente uma sequência $\{z^n\}$ (com $z^0 = z$) de Cauchy, cujo limite seja o ponto z_λ desejado. Seja z^n conhecido. Então, defina

   $$S_{n+1} := \{x \in X : x \neq z^n \text{ e } f(x) + \frac{\varepsilon}{\lambda} d(x, z^n) \leq f(z^n)\}.$$

   Se $S_{n+1} = \emptyset$, faça $z^{n+1} := z^n$. Caso contrário, escolha $z^{n+1} \in S_{n+1}$ tal que

   $$f(z^{n+1}) < \inf_{x \in S_{n+1}} f(x) + \frac{1}{2}(f(z^n) - \inf_{x \in S_{n+1}} f(x)).$$

   (a) Mostre que é possível, encontrar z^{n+1} em ambos casos.
   (b) Mostre que existe $\sum_{n=0}^{\infty} d(z^{n+1}, z^n)$ e como consequência que $\{z^n\}$ é uma sequência de Cauchy.
   (c) Prove que se $S_n = \emptyset$ para algum $n \in \mathbb{R}$, então $S_m = \emptyset$ para todo $m \geq n$, e assim $z_\lambda := z^n$ satisfaz as condições requeridas.

   Suponha, a partir de agora que $S_n \neq \emptyset$, para todo $n \in \mathbb{N}$ e denote por z_λ o limite de $\{z^n\}$

   (a) Mostre que $S_{n+1} \subset S_n$ para todo $n \in \mathbb{N}$ e que $z_\lambda \in \cap_{n \in \mathbb{N}} S_n$ (aqui usamos que f é lsc).
   (b) Verifique a desigualdade $f(z^{n+1}) - \inf\{f(x) : x \in S_{n+1}\} \leq f(z^n) - f(z^{n+1})$, prove que $f(z^n) - f(z^{n+1}) \to 0$, e use ditas propriedades para provar que $\cap_{n \in \mathbb{N}} S_n := z_\lambda$.
   (c) Use que $\cap_{n \in \mathbb{N}} S_n := z_\lambda$, para mostrar que se $y \neq z_\lambda$ (i.e. $y \notin \cap_{n \in \mathbb{N}} S_n$), temos que $f(z_\lambda) < f(y) + \frac{\varepsilon}{\lambda} d(z_\lambda, y)$. Complete a prova.

9. *Existência de pontos aproximadamente estacionários.* Seja $f : \mathbb{R}^m \to \mathbb{R}$ uma função derivável. Suponha que existe $x \in \mathbb{R}$ tal que $f(x) < \inf f + \varepsilon$ para certo $\varepsilon > 0$. Mostre que para todo $\lambda > 0$, existe um x^* com $\|x^* - x\| \leq \lambda$ e $\|\nabla f(x^*)\| \leq \varepsilon / \lambda$. O ponto x^* é chamado de ponto aproximadamente estacionário.
   Ainda mais, use o resultado anterior para provar que se $\inf f > -\infty$ e f derivável, existe uma sequência $\{x^n\}$ tal que $f(x^n) \to \inf f$ e $\|\nabla f(x^n)\| \to 0$.

10. *Teorema de Ponto Fixo de Caristi.* Seja (X, d) um e. m. completo e $\Phi : (X, d) \to \mathbb{R} \cup \{+\infty\}$ uma função lsc com limitada inferiormente. Se $T : X \rightrightarrows X$ é uma multifunção (i.e. $T(x)$ é um subconjunto de X) tal que

$$\Phi(y) \leq \Phi(x) - d(x, y) \quad \forall x \in X, y \in T(x).$$

Então, existe um $x^* \in X$ tal que $x^* \in T(x^*)$. *Dica: Use o principio variacional de Ekeland.*

11. *Teorema do ponto fixo de Banach.* Seja (X, d) e. m. completo e $T : X \to X$ uma contração, isto é, existe $\eta \in [0, 1)$ tal que $d(T(x), T(y)) \leq \eta d(x, y)$ para todo $x, y \in X$. Então, $x^* \in X$ tal que $x^* = T(x^*)$. *Dica:* Use o teorema de ponto fixo de Caristi com $\Phi(x) := (1 - \eta)^{-1} d(x, T(x))$.

12. *Condição de Palais-Smale* Seja $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1. Dizemos que f satisfaz a condição de Palais-Smale se para toda sequência $\{x^k\}$ tal que

$$f(x^k) \text{ converge para algum número } \alpha \quad \text{e} \quad \nabla f(x^k) \to 0,$$

a sequência $\{x^k\}$ tem uma subsequência convergente.

Prove que se f é limitada inferiormente e satisfaz a condição de Palais-Smale. Então, f deve ser coerciva.

13. Seja X um espaço vetorial real.

(a) Mostre que δ_C é convexo se, e somente se C é convexo.

(b) Mostre que $\operatorname{epi}(f)$ é convexo se, e somente se $f : X \to \mathbb{R} \cup \{+\infty\}$ é convexo.

(c) Mostre que $\sup\{f_i : i \in I\}$ é convexo se $f_i : X \to \mathbb{R} \cup \{+\infty\}$, $\forall i \in I$. Dê um exemplo de que $\inf\{f_i : i \in I\}$ não é necessáriamente convexo, mesmo se I fosse finito.

(d) Se $f : X \to \mathbb{R} \cup \{+\infty\}$ é convexo. Prove que $\operatorname{lev}_\gamma(f)$ é convexo para todo $\gamma \in \mathbb{R}$. Dê um exemplo onde a implicação reversa não vale.

(e) Mostre que f é convexa se, e somente se $f(\sum_{i=1}^n \alpha_i x_i) \leq \sum_{i=1}^n \alpha_i f(x_i)$, para todo $x_i \in \operatorname{dom}(f)$, $\alpha_i \geq 0$ e $\sum_{i=1}^n \alpha_i = 1$.

14. (slope inequality) Seja f uma função com $dom(f) = \mathbb{R}$. Então, f é convexa em $[a, b]$ se, e somente se para todos $x_0 < x < x_1$ em $[a, b]$, temos que

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_1) - f(x)}{x_1 - x}.$$

Faça um esboço dessas desigualdades.

15. *(Testes de convexidade usando derivadas).* Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável. Então, f é convexa se, e somente se algumas das siguentes condições valem:

- $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$, $\forall x, y$
- $\nabla f(x)$ é mononota. i.e. $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$, $\forall x, y$
- $\nabla^2 f(x) \geq 0$, para todo x. (aqui assumimos que f é duas vezes diferenciável)

Com essas ferramentas, facilmente podemos provar que $-\log(x)$, $\exp(x)$, (a entropia) $S(x) := -x \log(x)$, se $1 \geq x \geq 0$; $S(0) = 0$, $f(x) := x^{-\alpha}$, $\alpha > 0$, $x \in (0, \infty)$ são funções convexas.

16. Use a convexidade $x \to |x|^p$ para provar a desigualdade das médias generalizada (generalized mean inequality), isto é,

$$(\sum_{i=1}^n \alpha_i x_i^p)^{1/p} \leq (\sum_{i=1}^n \alpha_i x_i^q)^{1/q}$$

para todo $\alpha_i \geq 0$, com $\sum_{i=1}^n \alpha_i = 1$, $x_i > 0$, $\forall i$ e $p \leq q$, com $|p| + |q| \neq 0$. Como caso particular, prove a desigualdade de Young: Para todo $x, y \geq 0$, $p, q \in (1, \infty)$, tal que $1/p + 1/q = 1$ temos que

$$xy \leq \frac{1}{p} x^p + \frac{1}{q} y^q.$$

17. Seja f uma função convexa com valores reais extendidos.

(a) Mostre que o conjunto das soluções é convexa.

(b) Mostre que todo minimizador local é, de fato, um minimizador global

18. Seja A um subconjunto de (X, d). Mostre que $d_A(x) := \inf\{d(x, a) : a \in A\}$ é uma função Lipschitziana com constante de Lipschiz igual a 1.

19. Seja X um espaço finito dimensional com produto interno $\langle \cdot, \cdot \rangle$. Considere $C \neq \emptyset$ um conjunto fechado e convexo.

(a) Seja $x \in X$. Mostre que o problema

$$\min\{\|x - c\| : c \in C\}$$

admite uma *única* solução, qual é denotada por $\text{proj}_C(x)$, a projeção de x sobre o conjunto C. O valor otimo é denotado por $d_C(x)$. Se escolhemos $x = 0$, $\text{proj}_C(0)$ representa o elemento de C com norma minima.

(b) A projeção está caracterizada por uma desigualdade variacional. Isto é, mostre que

$$x^* := \text{proj}_C(x) \text{ se, e somente se } \langle x - x^*, c - x^* \rangle \leq 0, \text{ para todo } c \in C.$$

Se C é um subespaço vetorial, a desigualdade anterior se reduz a dizer que $x - x^* \perp C$.

(c) Em general, pode não existir elemento de norma minima. Seja $X = C[0, 1]$ o conjunto das função continua com a norma do supremo. Seja C o conjunto das funções $f \in X$ tal que

$$\int_0^{1/2} f(t)dt - \int_{1/2}^1 f(t)dt = 1.$$

Mostre que C é um conjunto não vazio, fechado e convexo, mas que não admite elemento de norma minima. (o problema aqui é que X não é reflexivo)