Problems for the mini-course

"Basics of optimization"

Kabatyansky G.A., Gasnikov A.V.

gasnikov@yandex.ru

Problem 1 (Lagrange multipliers principle and Implicit function theorem, 1.0). We have a sufficiently smooth optimization problem

$$f(x,y) \to \min_{g(x,y)=0}, \tag{1}$$

where $g:\mathbb{R}^{n+m} \to \mathbb{R}^m$. Assume that implicit function theorem can be applied for $g(x,y)=0$. That is, there exists smooth $y(x)$ such that $g(x,y(x)) \equiv 0$. Show that if (x_*,y_*) is a solution of problem (1) than

$$\exists\, \lambda : L_x(x_*,y_*,\lambda)=0, L_y(x_*,y_*,\lambda)=0,$$

where

$$L(x,y,\lambda)= f(x,y)+\langle \lambda, g(x,y)\rangle.$$

Literature

http://www.stat.uchicago.edu/~lekheng/courses/280/multipliers.pdf

Problem 2 (Danskin's formula, 0.5). Let $G(x,y)$ and

$$f(x)=\max_y G(x,y)\left(f(x)=\min_y G(x,y)\right)$$

are smooth enough functions. Assume that there exists $y(x)$ such that

$$G(x,y(x))=\max_y G(x,y)\left(G(x,y(x))=\min_y G(x,y)\right).$$

Than

$$\nabla f(x) = \nabla_x G(x, y(x)) = \left\{ \frac{\partial G(x,y)}{\partial x_i} \right\}_i \Bigg|_{y=y(x)} .$$

Literature

Bertsekas D.P. Nonlinear Programming. Belmont. MA: Athena Scientific, 1999.

https://arxiv.org/ftp/arxiv/papers/1506/1506.00292.pdf

Problem 3 (sensitivity in optimization, 1.0). We have a sufficiently smooth optimization problem

$$f(x) \to \min_{g(x)=b}, \tag{2}$$

where $g : \mathbb{R}^n \to \mathbb{R}^m$ ($m < n$). Let

$$L(x,\lambda) = f(x) + \langle \lambda, b - g(x) \rangle .$$

According to the Problem 1 (assume that all conditions we need we have) for optimal solution $x_*(b)$ of (2) there exists such $\lambda(b)$ that

$$L_x(x_*(b), \lambda(b)) = 0 .$$

Show that

$$\nabla F(b) = \lambda(b)$$

where

$$F(b) = \min_{g(x)=b} f(x) .$$

Literature

http://ocw.nctu.edu.tw/course/np982/Lecture11.pdf

Problem 4 (Lagrange multipliers principle and Separation theorem, 4.0). We have a convex optimization problem ($g : \mathbb{R}^n \to \mathbb{R}^m$)

$$f(x) \to \min_{\substack{g(x) \le 0 \\ x \in Q}} . \tag{3}$$

i) Show that the set $G = \{(g(x), f(x)), x \in Q\} \oplus \mathbb{R}_+^{m+1}$ is close convex set in the space of pairs (g, f).

ii) Let $f + g\lambda = \varphi(\lambda)$, where $\lambda \ge 0$ – be a tangent hyperplane to G. Show that

$$\varphi(\lambda) = \min_{x \in Q} \{ f(x) + \langle \lambda, g(x) \rangle \} \text{ (dual functional).}$$

iii) Obtain Lagrange multipliers principle from the Separation theorem for the set G and hyperplane from ii).

iv) Show that for arbitrary $x \in Q$, $g(x) \le 0$, $\lambda \ge 0$

$$f(x) \ge \varphi(\lambda)$$

$$\underbrace{\min_{x \in Q} f(x)}_{\substack{primal \\ problem}} \ge \underbrace{\max_{\lambda} \varphi(\lambda)}_{\substack{dual \\ problem}} . \text{ (weak duality)}$$

Give a geometric interpretation (it terms of i) – iii)) of this result.

v) Generalize iv) on general (not convex) case.

vi) Using iv) show that for the problem ($A \nsucc 0$)

$$f(x) = \langle x, Ax \rangle \to \min_{\|x\|_2^2 \le 1}$$

we have the following dual problem

$$\varphi(\lambda) = \begin{cases} -\lambda, \; A + \lambda I \succ 0 \\ -\infty, \text{ otherwise} \end{cases} \to \max_{\lambda} .$$

Show that there is no duality gap in this situation. That is we have strong duality

$$\min_{x \in Q} f(x) = \max_{\lambda} \varphi(\lambda)$$

vii) Consider NP-hard two-way partitioning problem

$$f(x) = \langle x, Wx \rangle \to \min_{x_i^2 = 1, \, i = 1, \ldots, n} .$$

Show that the dual problem is

$$\varphi(\lambda)=\begin{cases}-\sum_{i=1}^{n}\lambda_i, & W+Diag\{\lambda_i\}\succ 0\\ -\infty, & \text{otherwise}\end{cases}\to\max_{\lambda}.$$

Using weak duality show that

$$\min_{x_i^2=1,\, i=1,\dots,n} f(x)\geq n\lambda_{\min}(W).$$

Literature

Boyd S., Vandenberghe L. Convex optimization. Cambridge University Press, 2004.

http://stanford.edu/~boyd/cvxbook/

Problem 5 (minimax theorem, 1.0). Using Sion–Kakutani minimax theorem (see the very end of the book A. Nemirovski) build the dual problem for the following ELP problem

$$f(x)=\sum_{i=1}^{n}x_i\ln x_i\to\min_{\substack{Ax=b\\ x\in S_n(1)}},$$

where

$$S_n(1)=\left\{x\in\mathbb{R}_+^n:\sum_{k=1}^{n}x_k=1\right\}.$$

Why it is worth to solve dual problem instead of primal one?

Literature

Nemirovski A. Lectures on modern convex optimization analysis, algorithms, and engineering applications. Philadelphia: SIAM, 2013.

http://www2.isye.gatech.edu/~nemirovs/Lect_ModConvOpt.pdf

https://arxiv.org/ftp/arxiv/papers/1410/1410.7719.pdf

Problem 6 (Slater's conditions, 1.0). We have a convex optimization problem (3). We introduce

$$Q_{\bar{\lambda}}=\left\{\lambda\in\mathbb{R}_+^m:\varphi(\lambda)\geq\varphi\left(\bar{\lambda}\right)\right\}.$$

Assume that Slater's condition is true: *there exists such* $\bar{x} \in Q$ *that* $g(\bar{x}) < 0$. Show that

$$\max_{\lambda \in Q_{\bar{\lambda}}} \|\lambda\|_1 = \max_{\lambda \in Q_{\bar{\lambda}}} \sum_{i=1}^{m} |\lambda_i| \le \frac{1}{\gamma}\left(f(\bar{x}) - \varphi(\bar{\lambda}) \right),\ \gamma = \min_{i=1,\dots,m} \{-g_i(\bar{x})\}.$$

Hint: $\varphi(\bar{\lambda}) \le \varphi(\lambda) = \min_{x \in Q} \left\{ f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \right\} \le f(\bar{x}) + \sum_{i=1}^{m} \lambda_i g_i(\bar{x})$.

Literature

Bertsekas D.P. Nonlinear Programming. Belmont. MA: Athena Scientific, 1999.

Problem 7 (Alternating theorem, 1.5). Let $\{f_i(x)\}_{i=1}^{n}$ – convex functions. Show that one and only one of the following statements is true:

- There exists such x that $f_i(x) < 0$, $i = 1,\dots,n$;
- There exists such $y = \{y_i\}_{i=1}^{n} \ge 0\ (y \ne 0)$ that $\sum_{i=1}^{n} y_i f_i(x) \ge 0$ for all x.

Hint: *If first alternative is not true than introduce*

$$Z = \left\{ z = \{z_i\}_{i=1}^{n} :\ \varphi_i(x) < z_i,\ i = 1,\dots,n \text{ is solvable in } x \right\}.$$

Show that Z *is convex set and* $Z \cap \mathbb{R}_{-}^{n} = \varnothing$. *Then use separation theorem.*

Literature

Nemirovski A. Lectures on modern convex optimization analysis, algorithms, and engineering applications. Philadelphia: SIAM, 2013.
http://www2.isye.gatech.edu/~nemirovs/Lect_ModConvOpt.pdf

Problem 8 (Convex function and theirs representations, 3). i) Let $G(x,y)$ – is convex function as a function of x for all $y \in Y$. Assume that problem $\max_{y \in Y} G(x,y)$ is solvable for all x. Than $f(x) = \max_{y \in Y} G(x,y)$ is convex. Moreover $\partial f(x)$ can be calculated analogously to Danskin's formula (see Problem 2).

ii) Let $G(x,y)$ – is convex function as a function of x and y on the convex set Q. Assume that problem $\min_{y:(x,y)\in Q} G(x,y)$ is solvable for all x. Than $f(x) = \min_{y:(x,y)\in Q} G(x,y)$ is convex. How can $\partial f(x)$ be calculated?

Literature

Boyd S., Vandenberghe L. Convex optimization. Cambridge University Press, 2004.

http://stanford.edu/~boyd/cvxbook/

Nemirovski A. Lectures on modern convex optimization analysis, algorithms, and engineering applications. Philadelphia: SIAM, 2013.

http://www2.isye.gatech.edu/~nemirovs/Lect_ModConvOpt.pdf

https://arxiv.org/ftp/arxiv/papers/1506/1506.00292.pdf

Problem 9 (Schur complement, 2). Using Problem 7 ii) with

$$G(x,y) = \langle x, Cx\rangle + \langle y, Ay\rangle + 2\langle Bx, y\rangle,$$

show that if $A \succ 0$ (strictly) and

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succ 0.$$

Than

$$C - B^T A^{-1} B \succ 0.$$

Literature

Boyd S., Vandenberghe L. Convex optimization. Cambridge University Press, 2004.

http://stanford.edu/~boyd/cvxbook/

Problem 10 (Convex functions, 2). Do the following functions to be convex?

i) Let $f(x) = \ln\left(\sum_{k=1}^{n} \exp(\langle a_k, x\rangle)\right)$;

ii) $f(x,y) = \dfrac{(x-y)^2}{1-\max\{x,y\}}$, $x, y < 1$;

iii) $f(X) = \ln\det X^{-1}$, $X \succ 0$.

Literature

Boyd S., Vandenberghe L. Convex optimization. Cambridge University Press, 2004.