Preclass 07: Convex Optimization

[SCS4049] Machine Learning and Data Science

Seongsik Park (s.park@dgu.edu)

Department of Artificial Intelligence, Dongguk University

Lagrange multiplier

Minimize $f(x, y) = x + y$ subject to the constraint $x^2 + y^2 = 1$, i.e.,

$$g(x, y) = x^2 + y^2 - 1 = 0 \tag{1}$$

Hence,

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda g(x, y) = x + y + \lambda(x^2 + y^2 - 1) \tag{2}$$

Gradient

$$\nabla_{x,y,\lambda}\mathcal{L}(x, y, \lambda) = (1 + 2\lambda x, 1 + 2\lambda y, x^2 + y^2 - 1) \tag{3}$$

and therefore,

$$\nabla_{x,y,\lambda}\mathcal{L}(x, y, \lambda) \iff \begin{cases} 1 + 2\lambda x = 0 \\ 1 + 2\lambda y = 0 \\ x^2 + y^2 - 1 = 0 \end{cases} \tag{4}$$

$$\nabla_{x,y,\lambda}\mathcal{L}(x,y,\lambda) \iff \begin{cases} 1+2\lambda x = 0 \\ 1+2\lambda y = 0 \\ x^2+y^2-1=0 \end{cases} \tag{5}$$

This yields

$$x = y = -\frac{1}{2\lambda}, \quad \lambda \neq 0 \tag{6}$$

$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} - 1 = 0 \tag{7}$$

So,

$$\lambda = \pm\frac{1}{\sqrt{2}} \tag{8}$$

which implies that the stationary points of $\mathcal{L}$ are

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \quad \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) \tag{9}$$

Optimization problem in standard form

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) && && (10)\\
\text{subject to} \quad & f_i(x) \leq 0, && i = 1, 2, ..., m && (11)\\
& h_i(x) = 0, && i = 1, 2, ..., p && (12)
\end{aligned}
$$

- $x \in \mathcal{R}^n$ is the optimization variable
- $f_0 : \mathcal{R}^n \rightarrow \mathcal{R}$ is the objective or cost function
- $f_i : \mathcal{R}^n \rightarrow \mathcal{R}, i = 1, 2, ..., m$ are the inequality constraint functions
- $h_i : \mathcal{R}^n \rightarrow \mathcal{R}$ are the equality constraint functions

Convex optimization problem

Standard form convex optimization problem

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) && (13)\\
\text{subject to} \quad & f_i(x) \leq 0, \qquad i = 1, 2, ..., m && (14)\\
& a_i^T x = b_i, \qquad i = 1, 2, ..., p && (15)
\end{aligned}
$$

- $f_0, f_1, ..., f_m$ are convex
- equality constraints are affine

Often written as

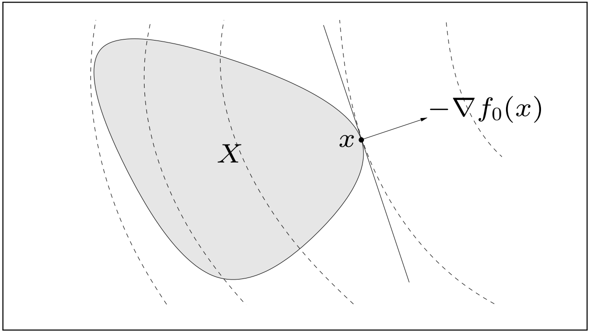
$$
\begin{aligned}
\text{minimize} \quad & f_0(x) && (16)\\
\text{subject to} \quad & f_i(x) \leq 0, \qquad i = 1, 2, ..., m && (17)\\
& Ax = b && (18)
\end{aligned}
$$

Important property: feasible set of a convex optimization problem is convex

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

$$\nabla f_0(x)^T(y - x) \geq 0 \quad \text{for all feasible } y \tag{19}$$



if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

Appendix

Reference and further reading

- "Chap 7 | Sparse Kernel Machines" of C. Bishop, Pattern Recognition and Machine Learning
- "Chap 5 | Support Vector Machines" of A. Geron, Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow
- "Chap 4 | Convex Optimization Problems", "Chap 5 | Duality" of S. Boyd, Convex Optimization