Lecture 14
Optimality conditions

– unconstrained & differentiable problems

– unconstrained & non-differentiable problems

– constrained & non-differentiable problems

– examples

Optimality conditions—unconstrained

$$\inf_{x\in\mathbf{R}^m} f(x)$$

assume f finite everywhere

x is said to minimize f over $\mathbf{R}^m$ if

$$\text{for every } y, \quad f(y) \geq f(x)$$

remember definition of subdifferential at x, $\partial f(x)$: set of g's such that

$$\text{for every } y, \quad f(y) \geq f(x) + g^T(x-y)$$

hence, **optimality condition:**

$$x \text{ is optimal} \iff 0 \in \partial f(x)$$

Example

piecewise linear minimization

$$f(x) = \sup_{1 \le i \le m} (a_i^T x + b_i)$$

subdifferential at x is

$$\partial f(x) = \mathbf{Co}\,\{a_i \mid i \text{ s.t. } a_i^T x^* + b_i = f(x)\}$$

then x^* minimizes f iff

$$0 \in \mathbf{Co}\,\{a_i \mid i \text{ s.t. } a_i^T x + b_i = f(x^*)\}$$

i.e., iff there exist $\lambda_i \ge 0$, $\Sigma_{i=1}^m \lambda_i = 1$, s.t.

$$\sum_{i=1}^m \lambda_i a_i = 0, \text{ and } \lambda_i = 0 \text{ if } a_i^T x^* + b_i < f(x^*)$$

the λ_i's are called Lagrange multipliers (non zero only at active constraints)

note *complementarity:* at optimum,

$$\text{for every } i, \quad \lambda_i(a_i^T x^* + b_i - f(x^*)) = 0$$

Differentiable, constrained problems

constrained minimization problem:

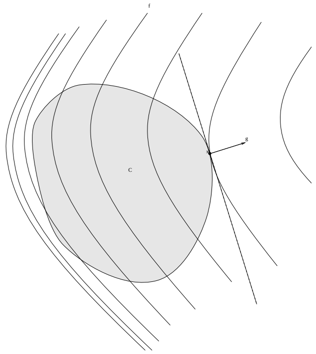
$$\inf_{x \in C} f_0(x)$$

where f is convex, **differentiable** (and C convex)

then $x \in C$ is optimal iff

$$y \in C \iff \nabla f_0(x)^T (y - x) \ge 0$$

- $-\nabla f_0(x)$ defines supporting hyperplane for C at x
- if you move from x towards any feasible y, f_0 does not decrease



- hence $x \in C$, $\nabla f_0(x) = 0$ implies x optimal
- for unconstr. problems, x is optimal iff $\nabla f_0(x) = 0$

Optimality conditions—constrained

consider

$$\textbf{inf}\ f_0(x) \text{ subject to } f_i(x) \le 0, \quad i = 1, \dots, m$$

where

- f_i convex and finite-valued (hence subdifferentiable)
- *Slater condition:* strict feasibility of above (primal) problem
- primal and dual optima are both attained

then a primal-dual pair x^*, λ^* is optimal iff

$$\boxed{\begin{aligned} f_i(x^*) &\le 0 \text{ (primal feasibility)} \\ \lambda_i &\ge 0 \text{ (dual feasibility)} \\ 0 &\in \partial f_0(x^*) + \sum_{i=1}^m \lambda_i \partial f_i(x^*) \text{ (stationarity)} \\ \lambda_i f_i(x^*) &= 0, \quad 1 \le i \le m \text{ (complementary slackness)} \end{aligned}}$$

these are the Karush-Kuhn-Tucker (KKT) conditions of the problem

Example

piecewise linear minimization

$$f(x) = \max_{1 \le i \le m} (a_i^T x + b_i)$$

in epigraph form:

$$\textbf{inf}\ t \text{ s.t. } a_i^T x + b_i \le t, \quad i = 1, \dots, m$$

apply previous result: pair (x^*, t^*) optimal iff

$$\begin{aligned} a_i^T x^* + b_i &\le t^* \quad 1 \le i \le m \\ \lambda_i &\ge 0, \quad 1 \le i \le m \\ \sum_{i=1}^m \lambda_i a_i &= 0 \\ \lambda_i (a_i^T x^* + b_i - t^*) &= 0, \quad 1 \le i \le m \end{aligned}$$

consistent with result obtained with unconstrained form of problem

Proof of necessity

Since strong duality holds, and primal and dual optima are both attained by some pair (x^*, λ^*), we have

$$\begin{aligned} f_0(x^*) &= g(\lambda^*) \text{ (zero optimal duality gap)} \\ &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \right) \text{ (definition of } g) \\ &\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) \text{ (by definition of } \mathbf{inf}) \end{aligned}$$

So:

$$\sum_{i=1}^m \lambda_i^* f_i(x^*) \geq 0$$

Since x^* is primal feasible, $f_i(x^*) \leq 0$, similarly, λ^* is dual feasible, so $\lambda^* \geq 0$; implies

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \ldots, m$$

above known as *complementary slackness* condition

Proof of necessity (follw'd)

from complementary slackness,

$$f_0(x^*) = g(\lambda^*) = L(x^*, \lambda^*)$$

so x^* is a minimizer of $L(x, \lambda^*)$ over x, so we must have

$$0 \in \partial_x L(x^*, \lambda^*)$$

(∂_x denotes the subdifferential wrt x)

above is equivalent to

$$0 \in \partial f_0(x^*) + \sum_{i=1}^m \lambda_i^* \partial f_i(x^*)$$

Proof of sufficiency

KKT conds are sufficient for primal-dual optimality for any convex opt problem (with finite-valued fcns)

If x^*, λ^* are any points that satisfy the KKT conditions

$$\begin{aligned} f_i(x^*) &\leq 0 \\ \lambda_i^* &\geq 0 \\ 0 &\in \partial f_0(x^*) + \Sigma_{i=1}^m \lambda_i^* \partial f_i(x^*) \\ \lambda_i^* f_i(x^*) &= 0, \end{aligned}$$

and f_i are convex, then x^* and λ^* are primal and dual optimal, with zero duality gap.

To see this, note that the third condition states that x^* minimizes $L(x, \lambda^*)$, thus

$$\begin{aligned} g(\lambda^*) &= L(x^*, \lambda^*) \\ &= f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) \\ &= f_0(x^*) \end{aligned}$$

where in the last line we use the fact that $\lambda_i^* f_i(x^*) = 0$.

This shows that x^* and λ^* have zero duality gap; hence they are primal and dual optimal.

Summary

For a convex optimization pb with finite-valued fcns:

- KKT conds. are *sufficient* for (primal and dual) optimality of a pair x, λ
- KKT conds. are *necessary* if strong duality holds and both primal and dual optima are attained
- If *both* primal and dual problems are strictly feasible, then strong duality holds and both primal and dual optima are attained

  In this case, KKT are nec. and suff. conds for optimality

Many algorithms can be interpreted as methods for solving the KKT equations

Optimality conditions for LP

(primal) LP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

optimal value $p^\star$

Lagrangian:

$$L(x, \lambda) = c^T x - \lambda^T (b - Ax)$$

where $\lambda \succeq 0$ (pay positive price if constraint violated)

dual form:

$$\begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0, \quad \lambda \succeq 0 \end{array}$$

optimal value $d^\star$

Fact: if either primal or dual is feasible, $p^\star = d^\star$.

KKT conditions for LP

primal-dual gap is

$$\begin{aligned} p - d &= c^T x + b^T \lambda \quad \text{(defn. of primal \& dual objectives)} \\ &= -(A^T y)^T x + b^T \lambda \quad (\text{since } c = -A^T y) \\ &= \lambda^T (b - Ax) \end{aligned}$$

zero duality gap with positive vectors x, $s = b - Ax$ implies complementarity:

$$\lambda_i (b - Ax)_i = 0, \quad i = 1, \ldots, m$$

KKT conditions:

$$\begin{gathered} s = b - Ax, \quad s \succeq 0, \\ c + A^T \lambda = 0, \quad \lambda \succeq 0, \\ \lambda_i s_i = 0, \quad i = 1, \ldots, m \end{gathered}$$

a feasible x solves LP if there is a dual feasible λ with zero duality gap

Optimality conditions for SDP

SDP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \cdots + x_n F_n \preceq F_0 \end{array}$$

Lagrangian (multiplier $Z = Z^T \in \mathbf{R}^{m \times m}$)

$$L(x, Z) = c^T x - \mathbf{Tr}\, Z F(x)$$

where $F(x) = F_0 - x_1 F_1 - \cdots - x_n F_n$

dual function

$$\begin{aligned} g(Z) &= \inf_x \left(c^T x + \mathbf{Tr}\, Z(x_1 F_1 + \cdots + x_n F_n - F_0) \right) \\ &= \begin{cases} -\mathbf{Tr}\, F_0 Z & \text{if } \mathbf{Tr}\, F_i Z + c_i = 0, \quad i = 1, \ldots, n \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

dual problem

$$\begin{array}{ll} \text{maximize} & -\mathbf{Tr}\, F_0 Z \\ \text{subject to} & \mathbf{Tr}\, F_i Z + c_i = 0, \quad i = 1, \ldots, n \\ & Z = Z^T \succeq 0 \end{array}$$

KKT conditions

primal-dual gap is

$$\begin{aligned} p - d &= c^T x + \mathbf{Tr}\, F_0 Z \quad \text{(defn. of primal \& dual objectives)} \\ &= \mathbf{Tr}\, F(x) Z \quad (\text{since } c_i = -\mathbf{Tr}\, F_i Z) \end{aligned}$$

zero duality gap with positive semidefinite matrices $F(x)$, Z implies **matrix** complementarity:

$$F(x) Z = 0$$

KKT conditions:

$$\begin{gathered} F(x) \succeq 0, \\ Z \succeq 0, \quad \mathbf{Tr}\, F_i Z + c_i = 0, \quad i = 1, \ldots, m \\ F(x) Z = 0 \end{gathered}$$

above are nec. and suff. if both primal and dual pbs are strictly feasible