Midterm 1

Kailong Wang

November 16, 2023

Q1: Normal cones to level sets. Grade:

Suppose $h : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable convex function, and consider the level set $L(0, h) = \{ x \in \mathbb{R}^n \mid h(x) < 0 \}$. Assuming that there exists some point $\bar{x} \in \mathbb{R}^n$ with $h(\bar{x}) < 0$, prove that, for any $x \in L(0, h)$, the normal cone $N_{L(0,h)}(x)$ to $L(0, h)$ at x is given by the formula

$$N_{L(0,h)}(x) = \{ \alpha \nabla h(x) \mid \alpha \geqslant 0, \alpha h(x) = 0 \} \begin{cases} \emptyset & \text{if } h(x) > 0 \\ \{ \alpha \nabla h(x) \mid \alpha \geqslant 0 \} & \text{if } h(x) = 0 \\ \{ \mathbf{0} \} & \text{if } h(x) < 0 \end{cases}$$

Solution

Proof. We prove the statement by cases.

1. If $h(x) > 0$, then $x \notin L(0, h)$, so $N_{L(0,h)}(x) = \emptyset$, which is trivial.
2. If $h(x) = 0$, x lies on the boundary of $L(0, h)$. The function h being continuously differentiable and convex implies that at x, the gradient $\nabla h(x)$ points in a direction that is normal to the level set $L(0, h)$ (since the normal cone is the polar cone of tangent cone of level set at x). By convexity, we have

$$h(y) \geqslant h(x) + \langle \nabla h(x), y - x \rangle, \qquad \forall y \in L(0, h).$$

   With $h(y) < 0$ (definition of level set) and $h(x) = 0$ (case assumption), we have

$$\langle \nabla h(x), y - x \rangle < 0.$$

   This means the vector $\nabla h(x)$ is an outward normal to the level set at x. Since h does not increase in the direction inside the level set, the normal cone at x consists of all non-negative scalar multiples of $\nabla h(x)$, *i.e.*

$$N_{L(0,h)}(x) = \{ \alpha \nabla h(x) \mid \alpha \geqslant 0 \}.$$

3. If $h(x) < 0$, the $x \in \operatorname{ri} L(0, h)$. The normal cone at a point in the relative interior of a convex set is just the zero vector. Because there are directions in every neighborhood around x that stay within $L(0, h)$, and therefore no "outside" direction is associated with a decrease from x within level set. Thus,

$$N_{L(0,h)}(x) = \{ \mathbf{0} \}.$$

□

Question: Is the necessity of the assumption $h(\bar{x}) < 0$ for some $\bar{x}$ to guarantee that the level set $L(0, h)$ is nonempty?

Q2 **Grade:**

(Optimality conditions for convex problems with "mixed" constraint sets.) Consider an optimization problem of the form

$$\begin{aligned} \min \quad & f(x) \\ \text{S.T.} \quad & Ax = b \\ & h_j(x) \leqslant 0 \qquad j = 1, 2, \ldots, r \\ & x \in X \end{aligned} \tag{1}$$

where

- $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a convex function
- A is an $m \times n$ matrix and $b \in \mathbb{R}^m$
- For $j = 1, 2, \ldots, r$, $h_j : \mathbb{R}^n \to \mathbb{R}$ is a differentiable convex function
- X is a convex set.

Let a_i denote row i of A, $i = 1, 2, \ldots, m$ represented as a column vector. Suppose that there exists a point $\bar{x} \in \mathbb{R}^n$ with the following properties:

- $\bar{x} \in \operatorname{ri} \operatorname{dom} f$
- $A\bar{x} = b$
- For $j = 1, 2, \ldots, r$, $h_j(\bar{x}) < 0$
- $\bar{x} \in \operatorname{ri} X$.

Show that for $x^* \in \mathbb{R}^n$ to be a solution of eq. (1), it is necessary and sufficient that there exist $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^r$ such that

$$\begin{aligned} \partial f(x^*) + \sum_{i=1}^{m} \lambda_i^* a_i + \sum_{j=1}^{r} \mu_j^* \nabla h_j(x^*) + N_X(x^*) &\ni 0 \\ \sum_{j=1}^{r} \mu_j^* h_j(x^*) &= 0 \\ Ax^* &= b \\ h_j(x^*) &\leqslant 0 \qquad j = 1, 2, \ldots, r \\ \mu_j^* &\geqslant 0 \qquad j = 1, 2, \ldots, r. \end{aligned}$$

Solution

Proof. Solving eq. (1) is equivalent to solving the following cvx problem:

1. $f_1(x) = f(x)$
2. $f_2(x) = \delta_L(x) = \begin{cases} 0, & Ax = b \\ +\infty, & \text{otherwise} \end{cases}$
3. $f_3(x) = \delta_X(x) = \begin{cases} 0, & x \in X \\ +\infty, & \text{otherwise} \end{cases}$

4. $f_4(x) = \delta_{h_j}(x) = \begin{cases} 0, & h_j(x) \leqslant 0 \\ +\infty, & \text{otherwise} \end{cases}$, j=1,2,...,r

Let x^* be an optimal solution to the optimization problem. For $h_j(x^*) = 0$, we can use the result from the previous question and have normal cone $N_{L(0,h_j)}(x^*) = \left\{ \mu_j^* \nabla h_j(x^*) \mid \mu_j^* \geqslant 0 \right\}$. For $h_j(x^*) < 0$, we will have $\mu_j^* = \mathbf{0}$ since we can't have a positive multiplier for a strictly feasible constraint. These contribute to the condition of $f_4(x)$

$$\sum_{j=1}^{r} \mu_j^* h_j(x^*) = 0; \qquad h_j(x^*) \leqslant 0 \quad j = 1, 2, \ldots, r; \qquad \mu_j^* \geqslant 0 \quad j = 1, 2, \ldots, r.$$

With the proposition that we have proved in the class, we know that solving $\partial(f_1 + f_2 + f_3)(x) \ni \partial f_1(x) + \partial f_2(x) + \partial f_3(x)$ with condition $Ax = b$ implies

$$\exists \lambda^* \in \mathbb{R}^m, x^* \in \mathbb{R}^n \qquad \partial f(x^*) + A^T\lambda^* + N_X(x^*) \ni 0.$$

Combine the above results (since $\bar{x} \in \operatorname{ri}\operatorname{dom} f_1 \cap \operatorname{ri}\operatorname{dom} f_2 \cap \operatorname{ri}\operatorname{dom} f_3 \cap \operatorname{ri}\operatorname{dom} f_4 \neq \emptyset$), we have (using Rockafellar-Moreau theorem)

$$\partial f(x^*) + \sum_{i=1}^{m} \lambda_i^* a_i + \sum_{j=1}^{r} \mu_j^* \nabla h_j(x^*) + N_X(x^*) \ni 0,$$

which completes the proof. □

Q3: Optimality conditions for convex cone programming. Grade:

Below, suppose $K \subseteq \mathbb{R}^m$ be a nonempty closed convex cone,

(a) Show that for any $x \in K$,
$$F_K(x) = \{\, z - \alpha x \mid z \in K, \alpha \geqslant 0 \,\}.$$

(b) Show that for any $x \in K$,
$$N_K(x) = \{\, y \in K^* \mid \langle x, y \rangle = 0 \,\}.$$
Hint: you may use the results of homework 3, problem 1(c) and 5(c).

(c) A is an $m \times n$ matrix, and $b \in \mathbb{R}^m$, and let $Z = \{\, x \in \mathbb{R}^n \mid Ax - b \in K \,\}$. Assume that Z is nonempty. Show that, for $x \in Z$,
$$N_Z(x) = \operatorname{cl}\left\{ A^T\lambda \mid \lambda \in K^*, \langle Ax - b, \lambda \rangle = 0 \right\}.$$
Hint: you may use the results of homework 3, problem 5.

(d) Show that $\operatorname{ri} Z \supseteq \{\, x \in \mathbb{R}^n \mid Ax - b \in \operatorname{ri} K \,\}$.

(e) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a convex function, suppose that the cone $A^T K^* = \left\{ A^T\lambda \mid \lambda \in K^* \right\}$ is closed, and consider the problem
$$\begin{aligned} \min \quad & f(x) \\ \text{S.T.} \quad & Ax - b \in K. \end{aligned} \tag{2}$$
Further suppose that there exists some point $\bar{x} \in \operatorname{ri}\operatorname{dom} f$ such that $A\bar{x} - b \in \operatorname{ri} K$. Show that, in order for $x^* \in \mathbb{R}^n$ to solve eq. (2), it is necessary and sufficient that there exists $\lambda^* \in \mathbb{R}^m$ such that
$$\begin{aligned} \partial f(x^*) + A^T\lambda^* &\ni 0 \\ \lambda^* &\in K^* \\ \langle Ax^* - b, \lambda^* \rangle &= 0. \end{aligned}$$

Solution

(a) *Proof.* By definition 4.6.1, given $x \in K$, we have

$$F_K(x) = \{\, y \in \mathbb{R}^m \mid x + \alpha y \in K, \forall \alpha \in [0, \bar{\alpha}], \bar{\alpha} > 0 \,\},$$

where y is a feasible direction. Here, to show that

$$F_K(x) = \{\, z - \alpha x \mid z \in K, \alpha \geqslant 0 \,\},$$

we need to show $y = z - \alpha x$ is a feasible direction and every feasible direction can be represented as $y = z - \alpha x$.

Let $y = z - \alpha x$ where $z \in K$ and $\alpha \geqslant 0$. Consider $x + ty$ for $t > 0$,

$$x + ty = x + t(z - \alpha x) = x + tz - t\alpha x = (1 - t\alpha)x + tz.$$

Since K is a convex cone, it is closed under positive linear combinations. For sufficiently samll t, $1 - t\alpha$ remains positive, and hence $(1 - t\alpha)x + tz$ is a positive linear combination of points in K, which means $x + ty \in K$ for all sufficiently small $t > 0$. Therefore, y is a feasible direction at x.

Let y be any feasible directions in $F_K(x)$. By definition, for all small $t > 0$, $x + ty \in K$. By the convexity of K, the line segment connecting x and $x + ty$ must entirely lie in K. For a sufficiently small t, this implies that y can be represented as

$$y = \frac{1}{t}(x + ty) - \frac{1}{t}x \Rightarrow (x + ty) - x = z - x,$$

where $z = x + ty \in K$. We can set $\alpha = 1$ to match the form required. So, $y = z - \alpha x$ with $x \in K$ and $\alpha = 1 \geqslant 0$. □

(b) *Proof.* The proof contains two parts.

(a) $N_K(x) \subseteq \{\, y \in K^* \mid \langle x, y \rangle = 0 \,\}$. Take any $y \in N_K(x)$. By definition of the normal cone, for all $z \in K$

$$\langle y, z - x \rangle \leqslant 0.$$

Since K is a cone, for $\lambda > 0$, λx is also in K. Replace z by λx and get

$$\langle y, \lambda x - x \rangle \leqslant 0,$$

which simplifies to

$$\lambda\langle y, x \rangle - \langle y, x \rangle \leqslant 0 \Rightarrow (\lambda - 1)\langle y, x \rangle \leqslant 0.$$

Since λ is arbitrary, we must have $\langle y, x \rangle = 0$. Besides, $\langle y, x \rangle \leqslant 0$ implies $y \in K^*$. Therefore, $N_K(x) \subseteq \{\, y \in K^* \mid \langle x, y \rangle = 0 \,\}$.

(b) $\{\, y \in K^* \mid \langle x, y \rangle = 0 \,\} \subseteq N_K(x)$. By the definition of K^*, we have that for every $y \in K^*$ and for every $z \in K$,

$$\langle y, z \rangle \leqslant 0.$$

And given $\langle y, x \rangle = 0$, we have

$$\langle y, z - x \rangle = \langle y, z \rangle - \langle y, x \rangle \leqslant 0 - 0 = 0,$$

which implies $y \in N_K(x)$. Therefore, $\{\, y \in K^* \mid \langle x, y \rangle = 0 \,\} \subseteq N_K(x)$.

The two parts together prove the statement. □

(c) *Proof.* For vector ν that is in $N_Z(x)$, it must have

$$\langle \nu, z - x\rangle \leqslant 0 \qquad \forall z \in Z.$$

To use the definition of Z, we have

$$\langle \nu, z - x\rangle = \langle A^T\nu, Az - Ax\rangle = \langle A^T\nu, (Az - b) - (Ax - b)\rangle.$$

For ν to be in the normal cone $N_Z(x)$, the last inner product should be non-positive for all $Az - b \in K$, which means that $A^T\nu$ must be in the normal cone to K at the point $Ax - b$, *i.e.* $A^T\nu \in N_K(Ax - b)$. With the previous question, we can related $N_K(Ax - b)$ to K^* by

$$N_K(Ax - b) = \{\, \lambda \in K^* \mid \langle Ax - b, \lambda\rangle = 0 \,\}.$$

So, a vector $\nu \in N_Z(x)$ must correspond to a λ in the dual cone K^* such that $A^T\lambda$ has a zero inner product with $Ax - b$, hence $\nu = A^T\lambda$ for some λ satisfying $\langle Ax - b, \lambda\rangle = 0$. To account for the fact that the normal cone $N_Z(x)$ is a closed set, we take the closure of the set $\left\{ A^T\lambda \mid \lambda \in K^*, \langle Ax - b, \lambda\rangle = 0 \right\}$ since the image under a linear transformation of a closed set is not necessarily closed. Therefore,

$$N_Z(x) = \operatorname{cl}\left\{ A^T\lambda \mid \lambda \in K^*, \langle Ax - b, \lambda\rangle = 0 \right\}.$$

□

(d) *Proof.* Let x_0 be such that $Ax_0 - b \in \operatorname{ri}(K)$. By the prolongation principle, for every point $\bar{y} \in K$, there exists $\delta > 1$ such that

$$Ax_0 - b + (\delta - 1)(Ax_0 - b - \bar{y}) \in K.$$

Now, let x be any point in Z, implying $Ax - b \in K$. Taking $\bar{y} = Ax - b$, the prolongation principle gives us

$$\begin{aligned} &Ax_0 - b + (\delta - 1)((Ax_0 - b) - (Ax - b)) \in K \\ &\Rightarrow A(x_0 + (\delta - 1)(x_0 - x)) - b \in K. \end{aligned}$$

showing that

$$x_0 + (\delta - 1)(x_0 - x) \in Z.$$

Since this is true for any $x \in Z$, x_0 must be in $\operatorname{ri}(Z)$.

We conclude that

$$\operatorname{ri} Z \supseteq \{x \in \mathbb{R}^n \mid Ax - b \in \operatorname{ri} K\}.$$

□

(e) *Proof.* Solving eq. (2) is equivalent to minimizing $f_1 + f_2$ where:

(a) $f_1(x) = f(x)$

(b) $f_2(x) = \delta_K(Ax - b) = \begin{cases} 0, & Ax - b \in K \\ +\infty, & \text{otherwise} \end{cases}$

The range of f_1 and f_2 are

(a) $\operatorname{ri}\operatorname{dom} f_1 = \operatorname{ri}\operatorname{dom} f$

(b) $\operatorname{ri}\operatorname{dom} f_2 = \operatorname{ri} Z$ (Z is in the same form in the previous question)

The condition says $\bar{x} \in \operatorname{ri} \operatorname{dom} f_1 \cap \operatorname{ri} K \neq \emptyset$. So $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$. Given x^* is optimal, we have

$$\begin{aligned}
& 0 \in \partial(f_1 + f_2)(x^*) \\
\Rightarrow\ & 0 \in \partial f_1(x^*) + \partial f_2(x^*) \\
\Rightarrow\ & 0 \in \partial f(x^*) + N_Z(x^*) \\
\Rightarrow\ & 0 \in \partial f(x^*) + A^T \lambda^* \qquad \text{for some } \lambda^* \in K^*
\end{aligned}$$

With $A^T\lambda^*$ is closed (from statement), combining $N_Z(x) = \operatorname{cl}\left\{ A^T\lambda \mid \lambda \in K^*, \langle Ax - b, \lambda \rangle = 0 \right\}$ from previous result, we finish the proof. $\square$