CMO: Constrained optimization for convex functions

Eklavya Sharma

1 Convex function and convex constraints

Let's analyze the following problem:

$$\begin{array}{lll} \min_x & f(x) & \\ \text{where} & c_i(x) \le 0 & \forall i \in I \\ & h_j(x) = 0 & \forall j \in I \end{array}$$

Here f and c_i are convex and C^1 and h_j is linear, i.e. $h_j(x) = a_j^T x - b_j$.

1.1 Feasible region is a convex set

Lemma 1 (Homework). *The set $S_i = \{x : c_i(x) \le 0\}$ is convex.*

Lemma 2 (Homework). *The set $S_j = \{x : h_j(x) = 0\}$ is convex.*

Lemma 3 (Homework). *The intersection of convex sets is convex.*

1.2 KKT point gives global minimum

Define the Lagrangian as

$$L(x, \lambda, \mu) = f(x) + \lambda^T c(x) + \mu^T h(x)$$

Lemma 4. *If $\lambda_i \ge 0$ and x is a feasible point, then $f(x) \ge L(x, \mu, \lambda)$.*

Proof. Since x is a feasible point,

$$\begin{aligned} & c_i(x) \le 0 \land h_j(x) = 0 \\ & \implies \lambda^T c(x) \le 0 \land \mu^T h(x) = 0 \\ & \implies f(x) + \lambda^T c(x) + \mu^T h(x) \le f(x) \\ & \implies L(x, \lambda, \mu) \le f(x) \end{aligned}$$

□

Lemma 5. *Let (x^*, λ^*, μ^*) be a KKT point. Then $f(x^*) = L(x^*, \mu^*, \lambda^*)$.*

Proof.

$$\begin{aligned}
&\lambda_i^* c_i(x^*) = 0 \wedge h_j(x^*) = 0 && \text{(complementary slackness and primal feasibility)}\\
&\implies \lambda^{*T} c(x^*) = 0 \wedge \mu^{*T} h(x^*) = 0\\
&\implies f(x^*) + \lambda^{*T} c(x^*) + \mu^{*T} h(x^*) = f(x^*)\\
&\implies L(x^*, \lambda^*, \mu^*) = f(x)
\end{aligned}$$

□

Theorem 6 (Proved previously)**.** *Let f be C^1 and convex. Then*

$$\forall u, v \in \mathbb{R}^d, f(v) \geq f(u) + \nabla_f(u)^T (v - u)$$

Theorem 7. *Let (x^*, λ^*, μ^*) be a KKT point. Then x^* is a constrained global minimum of f.*

Proof. Let x be a feasible point.

$$\begin{aligned}
f(x) &\geq L(x, \lambda^*, \mu^*) && \text{(by lemma 4)}\\
&= f(x) + \sum_i \lambda_i^* c_i(x) + \sum_j \mu_j^* (a_j^T x - b_j)\\
&\geq \left(f(x) + \nabla_f(x^*)^T (x - x^*)\right)\\
&\quad + \sum_i \lambda_i^* \left(c_i(x^*) + \nabla_{c_i}(x^*)^T (x - x^*)\right)\\
&\quad + \sum_j \mu_j^* \left(a_j^T (x - x^*) + (a_j^T x^* - b_j)\right) && \text{(by theorem 6)}\\
&= \left(f(x^*) + \sum_i \lambda_i^* c_i(x^*) + \sum_j \mu_j^* (a_j^T x^* - b_j)\right)\\
&\quad + (x - x^*)^T \left(\nabla_f(x^*) + \sum_i \lambda_i^* \nabla_{c_i}(x^*) + \sum_j \mu_j^* a_j\right) && \text{(rearrange terms)}\\
&= L(x^*, \lambda^*, \mu^*) + (x - x^*)^T (\nabla_x L)(x^*, \lambda^*, \mu^*)\\
&= f(x^*) && \text{(by lemma 5 and stationarity)}
\end{aligned}$$

Since for all feasible points $f(x) \geq f(x^*)$, x^* is a constrained global minimum of f. □

Note that unlike the necessary conditions for local minimum, here we do not require regularity.

1.3 Example: Projection over ball

Consider the optimization problem:

$$\min_x \frac{1}{2} \|x - z\|^2 \text{ where } \|x\|^2 \leq r^2$$

Here z lies outside the feasible region.

$\|x-z\|^2$ and $\|x\|^2$ are convex functions (because their hessian is $2I$, which is positive definite), so this is a convex optimization problem.

$$L(x, \lambda) = \frac{1}{2}\|x-z\|^2 + \lambda(\|x\|^2 - r^2)$$

Applying the KKT conditions, we get

- Stationarity: $z = (2\lambda + 1)x$.
- Primal feasibility: $\|x\|^2 \leq r^2$.
- Dual feasibility: $\lambda \geq 0$.
- Complementary slackness: $\lambda(\|x\|^2 - r^2) = 0$.

If we take $\lambda = 0$, then stationarity gives us $x = z$. This violates feasibility, so this is not possible. Therefore, complementary slackness gives us $\|x\|^2 = r^2$. On simplifying, we get

$$x = \frac{r}{\|z\|}z \qquad \lambda = \frac{1}{2}\left(\frac{\|z\|}{r} - 1\right) \qquad f(x) = \frac{1}{2}(r - \|z\|)^2$$