Convex Optimization - AI2101
Assignment - V

Bhuvan Chandra K
AI23BTECH11013

April 1, 2025

Question 3.29

A convex function $f : \mathbb{R}^n \to \mathbb{R}$, with $\operatorname{dom} f = \mathbb{R}^n$, is called piecewise-linear if there exists a partition of $\mathbb{R}^n$ as

$$\mathbb{R}^n = X_1 \cup X_2 \cup \cdots \cup X_L,$$

where $\operatorname{int} X_i \neq \emptyset$ and $\operatorname{int} X_i \cap \operatorname{int} X_j = \emptyset$ for $i \neq j$, and a family of affine functions

$$a_1^T x + b_1, \ldots, a_L^T x + b_L$$

such that

$$f(x) = a_i^T x + b_i \quad \text{for } x \in X_i.$$

Show that such a function has the form

$$f(x) = \max\{a_1^T x + b_1, \ldots, a_L^T x + b_L\}.$$

Solution:

f is a convex function, $f : \mathbb{R}^n \to \mathbb{R}$.
$\mathbb{R}^n$ is partitioned into regions $X_1, X_2, \ldots, X_l$, where X_i has a non-empty interior and X_i, X_j don't overlap if $i \neq j$.
On each region X_i, f is the affine function:

$$f(x) = a_i^T x + b_i.$$

Let:

$$g(x) = \max(a_1^T x + b_1, \ldots, a_l^T x + b_l).$$

We need to show $f(x) = g(x)$:

(i) $f(x) \leq g(x)$, $\forall x$. Let x^* be a point from $\mathbb{R}^n$. x^* must belong to one of the regions X_k.

$$f(x^*) = a_k^T x^* + b_k.$$

Since $g(x)$ is the maximum of all these affine functions:

$$g(x) = \max(a_1^T x + b_1, \ldots, a_l^T x + b_l) \geq a_k^T x + b_k \quad \forall x$$

Thus:

$$g(x) \geq f(x) \quad \forall x.$$

(ii) $f(x) \geq g(x)$, $\forall x$. Expressing from the previous case, we have to show:

$$a_k^T x^* + b_k \geq a_j^T x^* + b_j, \quad \forall j$$

We consider the contradiction:

Let there be a $j^* \neq k$ such that the inequality fails: i.e.,

$$a_{j^*}^T x + b_{j^*} > a_k^T x + b_k, \quad \text{for some } x \in X_k.$$

Since: $\text{int}(X_j) \neq \emptyset$, X_j has something in it.
Let $y \in \text{int}(X_j)$, $f(y) = a_j^T y + b_j$. Given f is convex, choose: $z = tx + (1-t)y$, where $t \in [0,1]$, on the line joining x and y.

By convexity:

$$f(z) \leq tf(x) + (1-t)f(y).$$

Note: As $t \to 0$, z goes into $\text{int}(X_j)$.
Thus:

$$f(z) = a_j^T z + b_j = a_j^T (tx + (1-t)y) + b_j,$$

$$f(z) = t(a_j^T x + b_j) + (1-t)(a_j^T y + b_j).$$

Note: Since on the line joining x, y, there should be some point such that $z \in X_k$. Let that point be z^*.

$$a_k^T z^* + b_k = a_j^T z^* + b_j.$$

i.e for points z from x to z^*

$$a_k^T z + b_k \geq a_j^T + b_j$$

Since x is on this line segment:

$$a_k^T x + b_k \geq a_j^T x + b_j.$$

This is a contradiction to our assumption. So our assumption is wrong.
So for any $x \in X_k$, we have:

$$a_k^T x + b_k \geq a_j^T x + b_j, \quad \forall j,$$

which gives:

$$f(x) \geq g(x).$$

From (i) and (ii):

$$f(x) = g(x) = \max(a_1^T x + b_1, \ldots, a_l^T x + b_l).$$

Question 3.30

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as

$$g(x) = \inf\{t \mid (x,t) \in \textbf{conv}\,\textbf{epi}\,f\}.$$

Geometrically, the epigraph of g is the convex hull of the epigraph of f.
Show that g is the largest convex underestimator of f. In other words, show that if h is convex and satisfies $h(x) \leq f(x)$ for all x, then $h(x) \leq g(x)$ for all x.

Solution:

We need to show that the convex hull or envelope of a function is the largest convex underestimator of that function.
For a function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\textbf{epi}\,f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : t \geq f(x)\}.$$

The convex hull of set S is the smallest convex set containing S, denoted by $\textbf{conv}\,S$.

As defined in the question, the envelope of $f(g(x))$ is:

$$g(x) = \inf\{t|(x,t) \in \mathrm{conv}(\mathrm{epi}\, f)\}.$$

(i) Let us prove $g(x)$ is convex.
The epigraph of g :

$$\mathrm{epi}\, g = \{(x,t)|g(x) \leq t\}.$$

Since $g(x)$ is defined using the infimum over $\mathrm{conv}(\mathrm{epi}\, f), g(x)$ is the smallest function whose epigraph equals $\mathrm{conv}(\mathrm{epi}\, f)$.

$$\mathrm{epi}\, g = \mathrm{conv}(\mathrm{epi}\, f).$$

By definition, $\mathrm{conv}(\mathrm{epi}\, f)$ is convex. So $g(x)$ has a convex epigraph, which makes $g(x)$ convex.

(ii) We show $g(x) \leq f(x)$, $\forall x$.
From the definition of $\mathrm{epi}\, f$: $(x, f(x)) \in \mathrm{epi}\, f$.
From the definition of the convex hull of $\mathrm{epi}\, f$:

$$\mathrm{epi}\, f \subseteq \mathrm{conv}(\mathrm{epi}\, f).$$

This means: $(x, f(x)) \in \mathrm{conv}(\mathrm{epi}\, f)$. Since $g(x)$ is the infimum of t for which $(x,t) \in \mathrm{conv}(\mathrm{epi}\, f)$, we get:

$$g(x) \leq f(x), \forall x.$$

(iii) Now we need to prove: If $k(x)$ is a convex function and $k(x) \leq f(x)$, $\forall x$, then $k(x) \leq g(x)$, $\forall x$.
From the assumption: $k(x)$ is a convex function and $k(x) \leq f(x)$, $\forall x$.
From the definition of the epigraph:

$$\mathrm{epi}\, k \supseteq \mathrm{epi}\, f.$$

Since $k(x)$ is lower than $f(x)$, it has more elements including that of $f(x)$, i.e., $\mathrm{epi}\, k$ contains $\mathrm{epi}\, f$.
As $k(x)$ is convex, $\mathrm{epi}\, k$ is also convex.
Note: $\mathrm{conv}(\mathrm{epi}\, f)$ is the smallest convex set containing $\mathrm{epi}\, f$. We get that since $\mathrm{epi}\, k$ is convex and contains $\mathrm{epi}\, f$, it also contains the convex hull of $\mathrm{epi}\, f$. Thus:

$$\mathrm{epi}\, k \supseteq \mathrm{conv}(\mathrm{epi}\, f).$$

So any point in $\mathrm{conv}(\mathrm{epi}\, f)$ should also be in $\mathrm{epi}\, k$.
If: $(x,t) \in \mathrm{conv}(\mathrm{epi}\, f)$, $t \geq k(x)$.
Let:

$$S_1 = \{t|\,(x,t) \in \mathrm{conv}(\mathrm{epi}\, f)\},$$
$$S_2 = \{t \mid t \geq h(x)\}.$$

As we have seen, $S_1 \subseteq S_2$.
From the result $S_1 \subseteq S_2 \implies \inf(S_1) \geq \inf(S_2)$. (For every element in S_1, $S_2's$ lower bound is also a lower bound.)
So:

$$g(x) = \inf\{t \mid (x,t) \in \mathrm{conv}(\mathrm{epi}\, f)\} \geq \inf\{t \mid t \geq h(x)\}.$$

Since:

$$\inf\{t \mid t \geq h(x)\} = h(x).$$

Hence:

$$g(x) \geq h(x), \quad \forall x$$

From (i), (ii), and (iii), g is the largest convex underestimator of f.

Question 3.31

Let f be a convex function. Define the function g as

$$g(x) = \inf_{\alpha>0} \frac{f(\alpha x)}{\alpha}.$$

(a) Show that g is homogeneous ($g(tx) = tg(x)$ for all $t \geq 0$). (b) Show that g is the largest homogeneous underestimator of f: If h is homogeneous and $h(x) \leq f(x)$ for all x, then we have $h(x) \leq g(x)$ for all x.
(c) Show that g is convex.

Solution:

Given f is a convex function:

$$g(x) = \inf_{\alpha>0} \frac{f(\alpha x)}{\alpha}$$

(a) For $t > 0$:

$$g(tx) = \inf_{\alpha>0} \frac{f(\alpha(tx))}{\alpha}.$$

Let us put $\beta = \alpha t$: If $\alpha > 0, t > 0$ then $\beta > 0$.
Thus:

$$g(tx) = \inf_{\beta>0} \frac{f(\beta x)}{\beta/t}.$$

Simplifying:

$$g(tx) = tg(x), \quad \forall t > 0.$$

So $g(x)$ is homogeneous.

(b) We need to show that $g(x)$ is the largest homogeneous underestimator of f.
If $h(x)$ is homogeneous and $h(x) \leq f(x)$, $\forall x$, then we have $h(x) \leq g(x)$, $\forall x$.
First, let's show that $g(x)$ is an underestimator of f. i.e,

$$g(x) \leq f(x), \ \forall x.$$

If we put $a = 1$ in the expression for $g(x)$:

$$g(x) = \inf_{\alpha>0} \frac{f(\alpha x)}{\alpha} \leq \frac{f(x)}{1}.$$

Since:

$$g(x) \leq f(x), \forall x,$$

g is an underestimator of f.
Let h be a homogeneous function. Assume $h(x) \leq f(x)$, $\forall x$. For any x, and any $\alpha > 0$:

$$h(\alpha x) = \alpha h(x) \leq f(\alpha x).$$

Dividing by α :

$$h(x) \leq \frac{f(\alpha x)}{\alpha} \quad \forall \alpha$$

Since the above equation holds for all α, the infimum over $\alpha > 0$ also holds:

$$h(x) \leq \inf_{\alpha>0} \frac{f(\alpha x)}{\alpha}$$

Thus:

$$h(x) \leq g(x), \forall x.$$

So g is the largest homogeneous underestimator of f.

(c) We need to show g is convex.

For any x_1, x_2, and any $\lambda \in [0,1]$:

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2).$$

Let $\epsilon > 0$, and let $\alpha_1, \alpha_2 > 0$.
From the definition of $g(x)$:

$$\frac{f(\alpha_1 x_1)}{\alpha_1} \leq g(x_1) + \epsilon,$$

and:

$$\frac{f(\alpha_2 x_2)}{\alpha_2} \leq g(x_2) + \epsilon.$$

Define: $z = \lambda d_1 x_1 + (1-\lambda)d_2 x_2$.
Since f is convex:

$$f(z) \leq \lambda \frac{f(\alpha_1 x_1)}{\alpha_1} + (1-\lambda)\frac{f(\alpha_2 x_2)}{\alpha_2}.$$

From the previous inequalities:

$$f(z) \leq \lambda(g(x_1) + \epsilon) + (1-\lambda)(g(x_2) + \epsilon)$$

$$f(z) \leq \lambda(g(x_1)) + (1-\lambda)(g(x_2)) + \epsilon$$

Since g is an underestimator of f,

$$g(z) \leq f(z).$$

Substituting:

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2) + \epsilon.$$

This holds for any $\epsilon > 0$.. So,

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2).$$

This shows that g is convex.

Question 3.33

Give a direct proof that the perspective function g, as defined in §3.2.6, of a convex function f is convex: Show that $\mathbf{dom}\, g$ is a convex set, and that for $(x,t), (y,s) \in \mathbf{dom}\, g$, and $0 \leq \theta \leq 1$, we have

$$g(\theta x + (1-\theta)y, \theta t + (1-\theta)s) \leq \theta g(x,t) + (1-\theta)g(y,s).$$

Solution:

Given $f : \mathbb{R} \to \mathbb{R}$ is a convex function. The perspective function $g : \mathbb{R}^{n+1} \to \mathbb{R}$ is defined as:

$$g(x,t) = tf\left(\frac{x}{t}\right), \quad t > 0.$$

We need to show that $\mathbf{dom}\, g$ is convex.

$$\mathbf{dom}\, g = \{(x,t) \mid \frac{x}{t} \in \mathbf{dom}\, f,\, t > 0\}.$$

Let:(x_1, t_1), $(x_2, t_2) \in \operatorname{dom} g$.
From the definition of $\operatorname{dom} g$:

$$\frac{x_1}{t_1} \in \operatorname{dom} f, \quad t_1 > 0,$$

and:

$$\frac{x_2}{t_2} \in \operatorname{dom} f, \quad t_2 > 0.$$

For some $\theta \in [0, 1]$,, Define:

$$(x_3, t_3) = (\theta x_1 + (1-\theta)x_2,\, \theta t_1 + (1-\theta)t_2).$$

We need to prove:

$$(x_3, t_3) \in \operatorname{dom} g.$$

From the definition:

$$t_3 = \theta t_1 + (1-\theta)t_2 > 0, (\text{all terms are positive}), \ \text{So} t_3 > 0.$$

Now we need:

$$\frac{x_3}{t_3} = \frac{\theta x_1 + (1-\theta)x_2}{\theta t_1 + (1-\theta)t_2} \in \operatorname{dom} f.$$

$$x_3/t_3 = \frac{\theta t_1}{\theta t_1 + (1-\theta)t_2} \cdot \frac{x_1}{t_1} + \frac{(1-\theta)t_2}{\theta t_1 + (1-\theta)t_2} \cdot \frac{x_2}{t_2}.$$

Define:

$$\beta = \frac{\theta t_1}{\theta t_1 + (1-\theta)t_2}, \quad (1-\beta) = \frac{(1-\theta)t_2}{\theta t_1 + (1-\theta)t_2}.$$

we have: $\beta > 0, \quad (1-\beta) > 0$.
Thus:

$$\frac{x_3}{t_3} = \beta\left(\frac{x_1}{t_1}\right) + (1-\beta)\left(\frac{x_2}{t_2}\right).$$

Since $x_1/t_1, x_2/t_2 \in \operatorname{dom} f$, and $\operatorname{dom} f$ is convex, we conclude:

$$\beta\left(\frac{x_1}{t_1}\right) + (1-\beta)\left(\frac{x_2}{t_2}\right) \in \operatorname{dom} f.$$

Thus:

$$x_3/t_3 \in \operatorname{dom} f.$$

We have $x_3/t_3 \in \operatorname{dom} f$, and $t_3 > 0$. So:$(x_3, t_3) \in \operatorname{dom} g$. Hence:

$$\operatorname{dom} g \text{ is convex.}$$

We need to show $g(x, t)$ is convex. Take $(x_1, t_1), (x_2, t_2) \in \operatorname{dom} g$, and let $\theta \in [0, 1]$. Define:

$$g(\theta(x_1, t_1) + (1-\theta)(x_2, t_2)) = g(\theta x_1 + (1-\theta)x_2,\, \theta t_1 + (1-\theta)t_2).$$

Expanding:

$$g(\theta x_1 + (1-\theta)x_2,\, \theta t_1 + (1-\theta)t_2) = (\theta t_1 + (1-\theta)t_2) f\left(\frac{\theta x_1 + (1-\theta)x_2}{\theta t_1 + (1-\theta)t_2}\right).$$

Like before, define:

$$\beta = \frac{\theta t_1}{\theta t_1 + (1-\theta)t_2}, \qquad (1-\beta) = \frac{(1-\theta)t_2}{\theta t_1 + (1-\theta)t_2},$$

Thus:

$$\frac{\theta x_1 + (1-\theta)x_2}{\theta t_1 + (1-\theta)t_2} = \frac{\beta x_1}{t_1} + \frac{(1-\beta)x_2}{t_2}$$

Since $f(x)$ is convex:

$$f\left(\beta\frac{x_1}{t_1} + (1-\beta)\frac{x_2}{t_2}\right) \leq \beta f\left(\frac{x_1}{t_1}\right) + (1-\beta) f\left(\frac{x_2}{t_2}\right).$$

Substituting back:

$$g(\theta x_1 + (1-\theta)x_2, \theta t_1 + (1-\theta)t_2) \leq (\theta t_1 + (1-\theta)t_2)\left[\beta f\left(\frac{x_1}{t_1}\right) + (1-\beta) f\left(\frac{x_2}{t_2}\right)\right].$$

Simplifying further:

$$g(\theta x_1 + (1-\theta)x_2, \theta t_1 + (1-\theta)t_2) \leq \frac{\theta t_1}{\theta t_1 + (1-\theta)t_2} g(x_1, t_1) + \frac{(1-\theta)t_2}{\theta t_1 + (1-\theta)t_2} g(x_2, t_2),$$

Thus:

$$g(\theta x_1 + (1-\theta)x_2, \theta t_1 + (1-\theta)t_2) \leq \theta g(x_1, t_1) + (1-\theta) g(x_2, t_2).$$

This shows that $g(x,t)$ is a convex function.