Convex Optimization
Assignment 3

1/8/2022
Jay Li

Chapter 3 Convex functions

Quasiconvex functions

3.42

Show W is quasiconcave.

Let $W(x) \geq \alpha$
$W(x') \geq \alpha$.

Now show $W(\theta x + (1-\theta)x') \geq \alpha$ for $0 \leq \theta \leq 1$

$$W(\theta x + (1-\theta)x') = \sup\{T \mid |\underbrace{\theta x^T f(t) + (1-\theta)x'^T f(t) - f_0(t)}_{\rho}| \leq \varepsilon \text{ for } 0 \leq t \leq T\}.$$

$$\rho = \theta(x^T f(t) - f_0(t)) + (1-\theta)(x'^T f(t) - f_0(t)).$$

since $|x^T f(t) - f_0(t)| \leq \varepsilon$ for $0 \leq t \leq \alpha$
$|x'^T f(t) - f_0(t)| \leq \varepsilon$ and since $0 \leq \theta \leq 1$

we have

$$\rho \leq \varepsilon \quad \text{for } 0 \leq t \leq \alpha.$$

$\Rightarrow W(\theta x + (1-\theta)) \geq \alpha$ for $0 \leq \theta \leq 1$.

i.e. the supersets are convex.

Easier to observe that

$W(x) \geq \alpha$ iff

$-\varepsilon \leq x_1 f(t) + \dots + x_n f(t) - f_0(t) \leq \varepsilon$

for all $t \in [0, \alpha)$

This is the intersection of an infinite number of half-spaces, i.e. a convex set

Log-concave and log-convex functions

3.54 $f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$

(a) verify $f''(x) f(x) \le f'(x)^2$ for $x \ge 0$.

$$f'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$f''(x) = -\frac{x}{\sqrt{2\pi}} e^{-x^2/2}$$

$$f'(x)^2 = \frac{1}{2\pi} e^{-x^2} > 0$$

And

$$f''(x) f(x) = -\frac{x}{2\pi} e^{-x^2/2} \int_{-\infty}^{x} e^{-t^2/2}\, dt \le 0$$

for $x \ge 0$

$\Rightarrow f''(x) f(x) \le f'(x)^2$ for $x \ge 0$ ▨

(b) verify $t^2/2 \ge -x^2/2 + xt$. for any t and x

$$t^2/2 + x^2/2 - xt = \left(\frac{t}{\sqrt{2}} - \frac{x}{\sqrt{2}}\right)^2 \ge 0$$

▨

Another way is to note $t^2/2$ is convex.

$$\Rightarrow t^2/2 \ge x^2/2 + x(t-x) = xt - x^2/2$$

According to $g(t) \ge g(x) + g'(x)(t-x)$ for differentiable convex function

(c) e^{-x} is a decreasing function of x.

$\Rightarrow$ ~~given~~ given $t^2/2 \geq -x^2/2 + xt$

$$e^{-t^2/2} \leq e^{x^2/2 - xt}$$

Taking the integral ~~of~~ of t from $-\infty$ to x keeps this inequality.

$$\Rightarrow \int_{-\infty}^{x} e^{-t^2/2}\, dt \leq e^{x^2/2} \int_{-\infty}^{x} e^{-xt}\, dt$$

$\square$

(d) ~~we~~ we now compute $\int_{-\infty}^{x} e^{-xt}\, dt$

$$= -\frac{1}{x} e^{-xt} \Big|_{-\infty}^{x}$$

$$= -\frac{1}{x} e^{-x^2} \quad \text{for} \quad x < 0.$$

$$\Rightarrow \int_{-\infty}^{x} e^{-t^2/2}\, dt \leq -\frac{1}{x} e^{\frac{x^2}{2}} e^{-x^2}$$

Multiply both sides by $-\frac{x}{2\pi} e^{-x^2/2} > 0$ for $x < 0$

we get

$$-\frac{x}{2\pi} e^{-x^2/2} \int_{-\infty}^{x} e^{-t^2/2}\, dt \leq \frac{1}{2\pi} e^{-x^2}$$

i.e. $f''(x) f(x) \leq f'(x)^2$ for $x < 0$

$\square$

Convexity with respect to generalized inequalities

3.57

Show that $f(X) = X^{-1}$ is matrix convex on S^n_{++}

Matrix convexity note:

$f: \mathbb{R}^n \to S^m$ is matrix convex iff

$$f(\theta x + (1-\theta)y) \preceq_{S^m_+} \theta f(x) + (1-\theta) f(y), \quad \theta \in [0,1]$$

According to Ex. 3.48 and we can show that this is equivalent to

$z^T f(x) z$ being convex for all z

i.e. we want to show

$z^T X^{-1} z$ being convex for all z, and $X \in S^n_{++}$.

This is proved in Example 3.4 by constructing its epigraph and noting that this epigraph is convex.

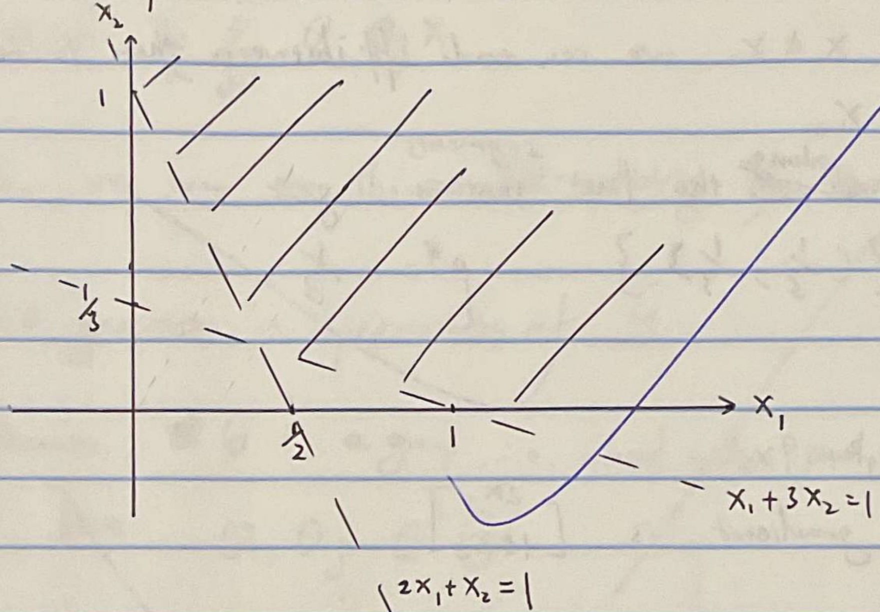
$\blacksquare$

Chapter 4 Convex Optimization Problems

Basic terminology and optimality conditions

4.1

sketch of the feasible set



(a) Because the objective is a linear function, ~~the optimal~~ one of the vertices must be an optimal point.

Solve $2x_1 + x_2 = 1$ and $x_1 + 3x_2 = 1$

get $x_1 = \frac{2}{5}$, $x_2 = \frac{1}{5}$, $\Rightarrow f_0 = \frac{3}{5}$

which is the optimal ~~point~~ value.

i.e. $X_{opt} = \{(\frac{2}{5}, \frac{1}{5})\}$ $p^* = \frac{3}{5}$.

(b) For, $f_0 = -x_1 - x_2$

we have $X_{opt} = \phi$ $p^* = -\infty$

(c) $X_{opt} = \{ (0, x_2) \mid x_2 \geq 1 \}$. $p^* = 0$

(d) $f_0 = \max\{x_1, x_2\}$.

If $x_1 \neq x_2$, we can trade off between them to reduce f_0.

$\Rightarrow x_1 = x_2$.

Solve ~~with~~ along the line ~~equations~~ segments, get

$X_{opt} = \{ (\frac{1}{3}, \frac{1}{3}) \}$. $p^* = \frac{1}{3}$

(e) $f_0 = x_1^2 + 9x_2^2$.

The gradient is $\begin{bmatrix} 2x_1 \\ 18x_2 \end{bmatrix}$.

Now, first search along the line $x_1 + 3x_2 \geq 1$.

At an optimal point, the gradient would be perpendicular to the line, if the optimal point is not an endpoint.

$\Rightarrow \begin{bmatrix} 2x_1 \\ 18x_2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 6x_1 - 18x_2 = 0$.

$\Rightarrow x_1 = 3x_2$, put into $x_1 + 3x_2 = 1$

$\Rightarrow x_1 = \frac{1}{2}$, $x_2 = \frac{1}{6}$.

It also satisfies $2x_1 + x_2 \geq 1$.

$\Rightarrow X_{opt} = \{ (\frac{1}{2}, \frac{1}{6}) \}$ $p^* = \frac{1}{2}$

4.4 (a)

$$\bar{x} = \frac{1}{k} \sum_{i=1}^{k} Q_i x$$

show that $\quad Q_i \bar{x} = \bar{x} \quad$ for $i = 1, \dots, k$.

$$Q_i \bar{x} = \frac{1}{k} \sum_{j=1}^{k} Q_i Q_j x$$

Now we want to show left-multiplying by Q_i on $G = \{Q_1, \dots, Q_k\}$

will generate a permutation of G.

Because G is a group, i.e. closed under products and inverse.

we have $\quad Q_i Q_j \in G$.

So let $\quad Q_i Q_j = Q_k$.

Now suppose there is Q_m where $Q_i Q_m = Q_k$ as well.

$$\Rightarrow \quad Q_i^{-1} Q_i Q_j = Q_i^{-1} Q_k$$

$$Q_i^{-1} Q_i Q_m = Q_i^{-1} Q_k$$

$$\Rightarrow \quad Q_j = Q_m = Q_i^{-1} Q_k.$$

Therefore no two Q_j gets mapped to same Q_k.

$$G \xrightarrow{Q_i} G$$

then because the number of element in left is the same as the number in the left. It is a one-to-one mapping

i.e. left-multiplying by Q_i generates a permutation.

for any i

$$\Rightarrow \sum_{j=1}^{k} Q_i Q_j x = \sum_{j=1}^{k} Q_j x$$

$$\Rightarrow \quad Q_i \bar{x} = \bar{x}$$

$$\Rightarrow \quad \bar{x} \in F.$$

(b) $f(\bar{x}) = f(\frac{1}{k} \Sigma Q_i x)$

$\le \frac{1}{k} \Sigma f(Q_i x)$ by convexity of f

$= \frac{1}{k} \Sigma f(x)$ by G-invariance

$= f(x)$

(c) let the optimal point be denoted x^*

we have $f_0(\bar{x^*}) \le f_0(x^*)$

by convexity and G-invariance of f_0.

Now we show $\bar{x}^* \in$ feasible set C.

$f_i(\bar{x}^*) \le \frac{1}{k} \sum_{j=1}^{k} f_i(Q_j x^*)$ $i = 1, \ldots, m$

And because $f_i(Q_j x^*) \le 0$, by G-invariance

$\Rightarrow f_i(\bar{x}^*) \le 0$.

i.e. $\bar{x}^* \in C$ the feasible set

$\Rightarrow$ $\bar{x}^*$ must be an optimal point.

because $\bar{x}^* \in F$

$\Rightarrow \exists$ optimal point in F

$\boxed{}$

(d) minimize $f(x)$. where $f(px) = f(x)$ for every permutation P

& $f(x)$ is convex.

by (c) we have that.

if $\exists$ a minimizer

then $\exists$ a minimizer $x \in F$

where $F = \{x \mid P.x = x, \text{ for all permutation } P\}$

$\Rightarrow F = \{\alpha \mathbb{1}\}$.

$\Rightarrow \exists$ x of the form $\alpha \mathbb{1}$.

$\boxed{}$

Linear optimization problems

4.8 (a)

$$\text{minimize} \quad c^T x$$
$$\text{subject to} \quad Ax = b$$

① If does not exist x st $Ax = b$

$\Rightarrow p^* = \infty$. and infeasible.

② Suppose $c \notin R(A^T)$

i.e. c can be decomposed as

$c = A^T\lambda + \hat{c}$ where $\hat{c} \neq 0$

and $\hat{c} \perp R(A^T)$

$\hat{c} \perp R(A^T)$ also means $A\hat{c} = 0$.

Now let x_0 be a solution to $Ax_0 = b$.

we can create $x_t = x_0 + t\hat{c}$

which also satisfies $Ax_t = b$

and $c^T x_t = (A^T\lambda + \hat{c})^T (x_0 + t\hat{c})$

$= \varphi + t\lambda^T A\hat{c} + t\hat{c}^T\hat{c}$

$=$ constant $+ 0 + t\|\hat{c}\|^2$

as $t \to -\infty$, $c^T x_t \to -\infty$.

$\Rightarrow p^* = -\infty$, optimal value not attained.

③ suppose $c \in R(A^T)$

i.e. $c = A^T\lambda$ for some λ.

let x be a solution to $Ax = b$

$\Rightarrow c^T x = \lambda^T A x = \lambda^T b$

i.e. $p^* = \lambda^T b$, and $X_{opt} = \{x \mid Ax = b\}$.

(b) minimize $c^T x$

subject to $a^T x \le b$ where $a \neq 0$

① Suppose $c^T a > 0$.

let x_0 be s.t. $a^T x_0 = b$.

let $x = x_0 - t a$ for $t > 0$.

$\Rightarrow a^T x = a^T x_0 - t a^T a \le b$

and $c^T x = c^T x_0 - t c^T a$

as $t \to \infty$, $c^T x \to -\infty$

$p^* = -\infty$, unattained.

② suppose $c = -\mu a$ for $\mu \ge 0$

$\Rightarrow c^T x = -\mu a^T x \ge -\mu b$

$p^* = -\mu b$, $X_{opt} = \{x \mid a^T x = b\}$.

③ c is not a ~~non-positive~~ ~~non-negative~~ ~~scale of~~ scaling of a.

Then $\exists\ \hat{\lambda}$ when $a^T\hat{\lambda} = 0$ & $c^T\hat{\lambda} < 0$.

let $x = x_0 + t\hat{\lambda}$ when $a^T x_0 = b$

$\Rightarrow \quad a^T x = a^T x_0 + t a^T \hat{\lambda} = b$

and $\quad c^T x = c^T x_0 + t\, c^T \hat{\lambda}$

as $t \to \infty$, $c^T x \to -\infty$.

i.e.

$p^* = -\infty$, unattained.

(c) minimize $c^T x$

subject to $l \preceq x \preceq u$.

where $l \preceq u$.

$$c^T x = \sum_{i=1}^{n} c_i x_i$$

$l_i \le x_i \le u_i$, $i = 1 \ldots, n$.

to minimize $c^T x$.

if $c_i > 0$, $x_i = l_i$

else if $c_i < 0$, $x_i = u_i$

else any x_i in range.

(d) minimize $c^T x$

subject to $\mathbb{1}^T x = 1$, $x \succeq 0$

minimize $\Sigma c_i x_i$

subject to $\Sigma x_i = 1$, $x_i \geq 0$

$p^* = \min c_i$ $\quad X_{opt} = \{ x \mid x_i = 1 \}$.

can allocate among the min c_i anyhow, if multiple min c_i

if $\mathbb{1}^T x \leq 1$ is the constraint.

① if all $c_i \geq 0$,

$p^* = 0$, $X_{opt} = \{0\}$

② Some $c_i < 0$

then put all x budget onto that c_i

i.e. $p^* = \min c_i$ $\quad X_{opt} = \{ x \mid x_i = 1 \}$.

again allocate anyhow if multiple min c_i

③ if $\min c_i = 0$

$p^* = 0$. $X_{opt} = \{ x \mid x_i = 0 \text{ for } c_i \neq 0 \}$.

(e) minimize $C^T X$

subject to $\mathbb{1}^T X = \alpha$, $0 \preceq X \preceq 1$

Reorder C_i from min to max.

$\{ C_{o_1} \cdots C_{o_n} \}$

Try to allocate from $\{ X_{o_1} \cdots X_{o_n} \}$

as much to the early X_i as possible

with the constraint $X_i \leq 1$

no difference if α is not integer

if $\mathbb{1}^T X \leq \alpha$ is the constraint.

Then only consider the negative C_is

$\{ C_{o_1} \cdots C_{o_k} \}$.

And allocate X_i only on these.

If $\min C_i = 0$.

then X_i for those i can be

anywhere from 0 to 1.

4.17 optimal activity levels

The constraints are

$$\sum_{j=1}^{n} A_{ij} x_j \leq c_i^{max} \quad , i=1,\ldots,m$$

$$x_j \geq 0 \quad , j=1,\ldots,n$$

The objective is

$$\text{minimize} \quad -\sum_{j=1}^{n} r_j(x_j)$$

which is convex, but not linear.

To transform into a LP
introduce variable t.

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & -\sum_{j=1}^{n} r_j(x_j) \leq t \\ & \sum_{j=1}^{n} A_{ij} x_j \leq c_i^{max} \quad , i=1,\ldots,m \\ & -x_j \leq 0 \quad , j=1,\ldots,n \end{array}$$

To transform $-\sum_{j=1}^{n} r_j(x_j) \leq t$ into a linear constraint.

we observe that this is equivalent to

2^n constraints where in each constraint $r_j(x_j)$ takes the form of

either $p_j x_j$

or $p_j q_j + p_j^{disc}(x_j - q_j)$.

~

Actually, what is better is to formulate as

$$\text{minimize} \quad \sum_{i=1}^{n} t_j$$

$$\text{subject to} \quad -r_j(x_j) \le t_j \quad , j = 1, \ldots, n$$

$$\sum_{j=1}^{n} A_{ij} x_j \le C_i^{max} \quad , i = 1, \ldots, m$$

$$-x_j \le 0 \quad , j = 1, \ldots, n$$

The $-r_j(x_j) = \max(-r_{j1}(x_j), -r_{j2}(x_j)) \le t_j$

can be written as

$$-r_{j1}(x_j) \le t_j$$
$$-r_{j2}(x_j) \le t_j$$

additional problems

1. Optimal activity levels (coding)

Solved problem both as a convex program and as a linear program.

Optimal activity levels : 4 , 22.5 , 31 , 1.5

revenues : 12 , 32.5 , 139 , 9

total revenue : 192.5

average prices : 3 , 1.444.. , 4.484 , 6.

prices are between p and p^{disc} as should.

3rd activity has high p and p^{disc} so makes sense to have ~~the~~ its activity level high.

2. Reformulating constraints in CVX.

(a) norm([x+2y , x-y]) = 0

The equality constraint can only take affine functions on both sides.

$\Rightarrow$ $x+2y=0$, $x-y=0$

(b) $[(x+y)^2]^2 \le x-y$

The square() is neither increasing nor decreasing. For it to be recognized as convex, its argument needs to be affine

$\Rightarrow$ $(x+y)^4 \le x-y$

(c) $1/x + 1/y \le 1$

$x \ge 0$

$y \ge 0$

$1/x$ is not convex over entire domain.

Therefore we use the special inv_pos() function that restricts the domain to be $\mathbb{R}_+$

$\Rightarrow$ inv_pos(x) + inv_pos(y) ≤ 1

(d) norm([max(x, 1) , max(y, 2)]) $\leq 3x+y$

norm() is convex, but neither increasing nor decreasing

it therefore need to take affine argument to be recognized as still convex.

$\Rightarrow$ use extra variable t_1, t_2.

norm ($[t_1, t_2]$) $\leq 3x+y$

$t_1 \geq$ max (x, 1)

$t_2 \geq$ max (y, 2)

(e) $xy \geq 1$

$x \geq 0$

$y \geq 0$

concave

xy is not a convex function

or geo_mean([x, y]) ≥ 1

$\Rightarrow$ x $\geq$ inv_pos(y)

or $\begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix}$ == semidefinite (2)

(f) $(x+y)^2$ / sqrt (y) $\leq x-y+5$

can not have division in DCP rule-set.

we can use quad_over_lin (a, b) function

which is a^2/b for $a \in \mathbb{R}$, and $b \in \mathbb{R}_+$

$\longrightarrow$

This is convex.

The second argument is deereasy, and can therefore take concave argument, and maintain convexity.

sqrt() is concave.

$\Rightarrow$ quad_over_lin(x+y, sqrt(y)) $\le$ x - y + 5

(g) $x^3 + y^3 \le 1$

$x \ge 0$

$y \ge 0$

x^3 is not convex for $x \le 0$

therefore is not acceped by DCP ruleset

$\Rightarrow$ pow_pos(x) + pow_pos(y) $\le$ 1.

(h) $x + z \le 1 + \text{sqrt}(xy - z^2)$

$x \ge 0$

$y \ge 0$

xy is not concave, so does not ~~go~~ have sqrt $(xy - z^2)$ reegnized as concave.

Instead, we can use geometric mean

$\sqrt{y(x - \frac{z^2}{y})}$ which is concave and increasy in args.

$\Rightarrow$ x+y $\le$ 1+ geo_mean (y, x - quad_over_lin(z, y))

$x \ge 0$

3. The illumination problem

Equal lamp power

p: [0.345 , 0.345, ...]

f: 0.469

Least squares w/ saturation

p: [1, 0, 1, 0, 0, 1, 0, 1, 0, 1]

f: 0.863

Regularized least squares

p: [0.499 , 0.477, 0.083, 0.00187, . . .

f: 0.444

Chebyshev approximation

p: [1 , 0.116 , 0 , 0, 1, 0, 1, 0.025,
0 , 1]

f: 0.4198

exact solution

p: [1 , 0.2023 , 0, 0, 1, 0, 1, 0.18816,
0, 1]

f: 0.3575