Convex Optimization Homework Solutions

Andres Espinosa

December 8, 2024

Contents

1 Homework 1

1.1 Exercise 2.1

Showing the definition of convexity for arbitrary k. $\theta_1 x_1 + \cdots + \theta_k x_k \in C$.

$$\theta_1 x_1 + \cdots + \theta_k x_k = \theta_1 x_1 + (1-\theta_1)(\mu_2 x_2 + \mu_3 x_3 + \cdots + \mu_k x_k) \tag{1}$$

Where each $\mu_n = \frac{\theta_n}{1-\theta_1}$

$$\sum_i \mu_i = \frac{\sum_i \theta_i}{1-\theta_1} = \frac{1-\theta_1}{1-\theta_1} = 1 \tag{2}$$

1.2 Exercise 2.2

The intersection between any two convex sets is convex. A line segment by definition is convex because it contains itself. Therefore, if there is any line segment L that intersects C, for C to be convex, the intersection must also be convex.

Similarly, for an affine set, the definition is that the set must contain hte line through any distinct points in the set. Therefore if there is a line that line segment L that intersects an affine set A, its intersection is also affine.

1.3 Exercise 2.5

Distance between two parallel hyperplanes $\{x \in \mathbb{R}^n | a^T x = b_1\}, \{x \in \mathbb{R}^n | a^T x = b_2\}$:

$$a^T(x_1 + at) = b_2 \tag{3}$$

$$a^T x_1 + a^T at = b_2 \tag{4}$$

$$t = \frac{b_2 - a^T x_1}{a^T a} \tag{5}$$

$$t = \frac{b_2 - b_1}{a^T a} \tag{6}$$

$$\frac{|b_1 - b_2|}{||a||_2} \tag{7}$$

1.4 Exercise 2.7

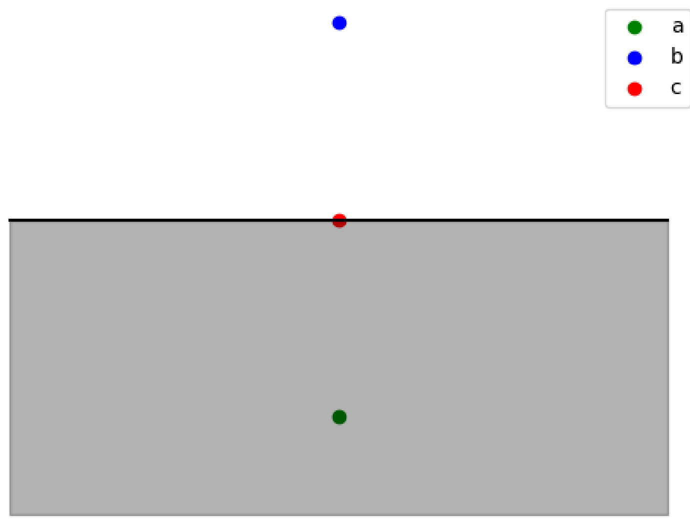
The image below illustrates the intuition behind this in $\mathbb{R}^2$. The set of all points that are closer to a than b $\{x | ||x-a||_2 \leq ||x-b||_2\}$, is a halfspace of the form

$$\{x | c^T x \leq d\} \tag{8}$$

where c is the normal vector that points from a to b, $c = \frac{a-b}{||a-b||_2}$ and $d = (\frac{a-b}{||a-b||_2})^T \frac{a+b}{2}$

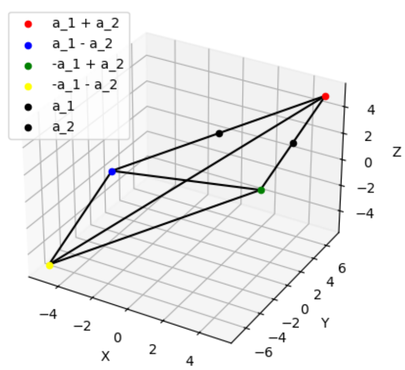
1.5 Exercise 2.8

Finding which of the sets are polyhedra



1.5.1 a

$S = \{y_1 a_1 + y_2 a_2 | -1 \leq y_1 \leq 1, -1 \leq y_2 \leq 1\}, a_1, a_2 \in \mathbb{R}^n$ This set is a 2D parallelogram in any $\mathbb{R}^n$. The image shown is one that is generated with $\mathbb{R}^3$ coordinates $a_1 = (1, 2, 3)$ and $a_2 = (4, 5, 2)$. It can be defined as the intersection between four halfspaces with their normals pointing from the origin to and centered on $a_1, a_2, -a_1, -a_2$.



The normal vectors for the halfspaces are:

$$A = \begin{bmatrix} \frac{a_1}{||a_1||_2} \\ \frac{a_2}{||a_2||_2} \\ \frac{-a_1}{||a_1||_2} \\ \frac{-a_2}{||a_2||_2} \end{bmatrix} \tag{9}$$

and the b vector is:

$$b = \begin{bmatrix} ||a_1||_2 \\ ||a_2||_2 \\ ||a_1||_2 \\ ||a_2||_2 \end{bmatrix} \tag{10}$$

Therefore, the set S can be expressed as $S = \{x|Ax \preceq b\}$

1.5.2 b

This is a polyhedra where:

$$A = I^n \tag{11}$$
$$b = 0 \tag{12}$$

$$C = \begin{bmatrix} \mathbf{1} \\ a \\ a^2 \end{bmatrix} \tag{13}$$

$$d = \begin{bmatrix} 1 \\ b_1 \\ b_2 \end{bmatrix} \tag{14}$$

1.5.3 c

Not attempted

1.5.4 d

Not attempted

1.6 Exercise 2.11

Showing that the hyperbolic set $\{x \in \mathbb{R}^2_{++}|x_1x_2 \geq 1\}$ is convex. As a generalization, show that $\{x \in \mathbb{R}^2_+|\Pi^n_{i=1}x_i \geq 1\}$

1.7 Exercise 2.12

Finding which of the following sets are convex:

1.7.1 a. slab

A set of the form $\{x \in \mathbb{R}^n|\alpha \leq a^Tx \leq \beta\}$, A slab is the intersection of two halfspaces $\mathbb{C}$ and $\mathbb{D}$ where $\mathbb{C} = \{x|a^Tx \leq \beta\}$ and $\mathbb{D} = \{x| - a^Tx \leq -\alpha\}$

1.7.2 b. rectangle

A set of the form $\{x \in \mathbb{R}^n|\alpha_i \leq x_i \leq \beta_i, i = 1, \ldots, n\}$. This is also the intersection of halfspaces, but this time it is the intersection of four halfspaces (in $\mathbb{R}^2$). But this is a polyhedra.

1.7.3 c. wedge

A set of the form $\{x \in \mathbb{R}^n | a_1^T x \leq b_1, a_2^T x \leq b_2\}$. Intersection of two halfspaces

1.7.4 d. point-set distance

A set of the form $\{x | \|x - x_0\|_2 \leq \|x - y\|_2 \forall y \in S\}$ where $S \subseteq \mathbb{R}^n$. This can be viewed as the intersection of many euclidean balls. The intersection of convex spaces is convex so the here, the set can be turned into $\cap_{y \in S} \{x | \|x - x_0\|_2 \leq \|x - y\|_2\}$

1.7.5 e. set-set distance

A set of the form $\{x | \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\}$ This set is not convex. Consider two crescent-moon shaped sets that lie next to each other so the tip of one moon is in the center of the other moon. The set that is closer to either of them is not convex because the crescent shape would stay in the set and is not convex.

1.7.6 f. affine set combo

A set of the form $\{x | x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbb{R}^n$ Yes, this is essentially the section of S_2 that can be translated to the convex S_1

1.7.7 g. fraction distance

Not attempted

1.8 Exercise 2.15

Finding which conditions sets are convex that are conditions on the probability simplex $\{p | \mathbf{1}^T p = 1, p \succeq 0\}$:

1.8.1 a. bounded expectation

This is a linear bound, so this set is convex.

1.8.2 b. prob a and b

Not attempted

2 Homework 2

2.1 Exercise 2.28

A matrix is positive semi-definite if we define two matrices symmetric A and Q, such that $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, Q and A are positive semidefinite if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x}$.
In other words, for a matrix A to be positive semidefinite, these two conditions must be satisfied:

- A is symmetric
- $x^T Ax \geq 0$ for all x

The second condition can be broadened to multiple equivalent definitions:

- $x^T Ax \geq 0$ for all x
- All eigenvalues are non-negative
- There exists a matrix B s.t. $B^T B = A$
- All principal minors are non-negative

2.1.1 n=1

For $n = 1$, the cone is defined by the inequality $x_1 \geq 0$

2.1.2 n=2

$x_1x_3 - x_2^2 \geq 0$ and $x_1, x_3 \geq 0$

2.1.3 n=3

All principal minors must be non-negative, so blocking off each row and column of the matrix.

Diagonals: $x_1, x_4, x_6 \geq 0$

Full matrix determinant: $x_1(x_4x_6 - x_5^2) - x_2(x_2x_6 - x_3x_5) + x_3(x_2x_5 - x_3x_4) \geq 0$ 2x2 determinants: $x_4x_6 - x_5^2 \geq 0$, $x_1x_6 - x_3^2 \geq 0$ $x_1x_4 - x_2^2 \geq 0$

2.2 Exercise 2.33

2.2.1 part a

A convex cone $K \subseteq \mathbb{R}^n$ is a proper cone if

- K is closed (contains its boundary)
- K is solid (nonempty interior)
- K is pointed (contains no line)

For this problem, we must show that the monotone non-negative cone defined as

$$K_{m+} = \{\mathbf{x} \in \mathbb{R}^n | x_1 \geq x_2 \geq \cdots \geq x_n \geq 0\} \tag{15}$$

is a proper cone.

Proving the first, where K_{m+} is closed and contains its boundary, we first find the boundary of K_{m+}. The boundaries are the two vectors in the direction of

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{16}$$

With magnitude x_1. These boundaries clearly satisfy the $\geq$ requirements of the set. If one of the conditions was a strict equality, then the convex cone would be open.

Proving the second, since all vectors are contained in between the boundaries, the set is solid.

Proving the third, because of the ≥ 0 requirement, the $\mathbf{x} \in \mathbb{R}^n$ of the set also have the $\mathbf{x} \succeq 0$ applicable. Therefore the cone is pointed.

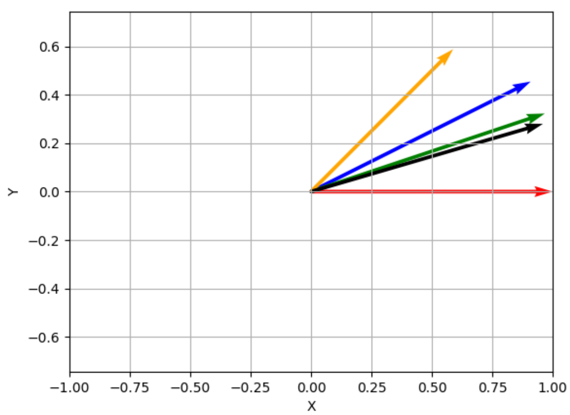


Figure 1: Image of a monotone non-negative cone in $\mathbb{R}^2$

2.2.2 part b

Finding the dual cone K^*_{m+}. A dual cone is defined as

$$K^* = \{y | x^T y \geq 0 \text{ for all } x \in K\} \tag{17}$$

$$x^T y \tag{18}$$

$$= \sum_{i=1}^{n} x_i y_i \tag{19}$$

$$= (x_1 - x_2)y_1 + (x_2 - x_3)(y_1 + y_2) + \cdots + (x_{n-1} - x_n)(y_1 + \cdots + y_n - 1) + x_n(y_1 + \cdots + y_n) \tag{20}$$

Since each term of y is getting multiplied by $x_i - x_{i+1}$, and $x_i \geq x_{i+1}$, each y term is getting multiplied by a non-negative x term, so if we want $x^T y \geq 0$ for all x, then we will need in the worst case scenario, for each one of the y terms to be positive.

$$y_1 \geq 0, y_1 + y_2 \geq 0, y_1 + y_2 + y_3 \geq 0, \ldots, \sum_{i=1}^{n} y_i \geq 0 \tag{21}$$

These conditions are all satisfied when $y_n \geq y_{n-1} \geq \cdots \geq y_2 \geq y_1 \geq 0$. Therefore, the dual cone is

$$K^* = \{y \in \mathbb{R}^n | y_n \geq y_{n-1} \geq \cdots \geq y_2 \geq y_1 \geq 0\} \tag{22}$$

2.3 Exercise 3.2

The first function is potentially quasiconvex because the sublevel sets shown are convex. The second function does not seem to have convex sublevel sets. So it is not quasiconvex or quasiconcave.

2.4 Exercise 3.5

Showing that the running average of a convex function is also convex. Since f is differentiable, we can attempt to differentiate F twice and evaluate it on the domain of F

$$F(x) = \frac{1}{x} \int_0^x f(t)dt \tag{23}$$

$$F'(x) = -\frac{1}{x^2} \int_0^x f(t)dt + \frac{1}{x} \frac{d}{dx} \int_0^x f(t)dt \tag{24}$$

$$= -\frac{1}{x^2} \int_0^x f(t)dt + \frac{1}{x}(f(x) + \int_0^x \frac{d}{dx} f(t)dt) \tag{25}$$

$$F'(x) = -\frac{1}{x^2} \int_0^x f(t)dt + \frac{f(x)}{x} \tag{26}$$

$$F''(x) = \frac{2}{x^3} \int_0^x f(t)dt - \frac{1}{x^2} \frac{d}{dx} \int_0^x f(t)dt + \frac{f'(x)x - f(x)}{x^2} \tag{27}$$

$$F''(x) = \frac{2}{x^3} \int_0^x f(t)dt - \frac{f(x)}{x^2} + \frac{f'(x)x - f(x)}{x^2} \tag{28}$$

$$F''(x) = \frac{2}{x^3} \int_0^x f(t)dt - \frac{2f(x)}{x^2} + \frac{f'(x)x}{x^2} \tag{29}$$

$$F''(x) = \frac{2}{x^3}(\int_0^x f(t)dt - f(x)x + \frac{f'(x)x^2}{2}) \tag{30}$$

$$= \frac{2}{x^3}(\int_0^x f(t) - f(x) - f'(x)(t - x)dt) \tag{31}$$

The term $f(t) - f(x) - f'(x)(t - x)$ is always positive because of the definition below for any convex function f

$$f(t) \geq f(x) + f'(x)(t - x) \tag{32}$$

Therefore, since the domain of F is confined to all positive numbers, the hessian of F is always greater than zero.

2.5 Exercise 3.6

- The epigraph of a function is a halfspace if the function is affine. Since the function is linear and infinite, its epigraph is a halfspace.
- If a function is linear then its epigraph is a convex cone.
- If a function is piecewise affine, then its epigraph is a polyhedron.

2.6 Exercise 3.15

2.6.1 part a

Showing that for $x > 0, u_0(x) = \lim_{\alpha \to 0} u_\alpha(x)$

$$u_\alpha(x) = \frac{x^\alpha - 1}{\alpha} \tag{33}$$

L'hopital :)

$$\lim \frac{x^\alpha - 1}{\alpha} = \lim x^\alpha \log x \tag{34}$$

This goes to $\log x$ as $\alpha \to 0$

2.6.2 Part b

Firstly, u_α is monotonically increasing because the first derivative is

$$u'_\alpha = x^{\alpha-1} \tag{35}$$

Which is always positive. The function is concave since the hessian is

$$u''_\alpha = (\alpha - 1)x^{\alpha-2} \tag{36}$$

Since $0 < \alpha \leq 1$, the function is affine for $\alpha = 1$ and concave for all other values of α. The equation $u_\alpha(1) = 0$ is also satisfied since $1^\alpha = 1 \forall \alpha$

2.7 Exercise 3.16

Finding whether the functions are convex, concave, quasiconcave, or quasiconvex.

2.7.1 Part b

$$f(x_1, x_2) = x_1 x_2 \tag{37}$$

$$\nabla^2 f = \begin{bmatrix} 01 \\ 10 \end{bmatrix} \tag{38}$$

In reducing to row echelon form, rows 1 and 2 would need to exchange and therefore the matrix is not positive semidefinite and not convex. It is not concave since the negative of the matrix is not positive semidefinite. It is quasiconcave but I didn't really know that.

2.7.2 Part c

$$f(x_1, x_2) = \frac{1}{x_1 x_2} \tag{39}$$

$$\nabla^2 f = \begin{bmatrix} \frac{2}{x_1^3 x_2} \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} \frac{2}{x_1 x_2^3} \end{bmatrix} \tag{40}$$

This function is convex if $\frac{2}{x_1^3 x_2} \geq 0$ and $\frac{3}{x_1^4 x_2^4} \geq 0$. Since $x_1, x_2 \in \mathbb{R}^2_{++}$, the function is convex.

2.7.3 Part d

$$f(x_1, x_2) = \frac{x_1}{x_2} \tag{41}$$

$$\nabla^2 f = \begin{bmatrix} 0 \frac{-1}{x^2} \\ \frac{-1}{x^2} \frac{2x_1}{x_2^3} \end{bmatrix} \tag{42}$$

This function is not convex or concave. I still don't really get quasiconvex or quasiconcave tbh.

2.7.4 Part e

$$f(x_1, x_2) = \frac{x_1^2}{x_2} \tag{43}$$

I know this is convex because I watched the lecture but here the hessian

$$\nabla^2 f = \begin{bmatrix} \frac{2}{x_2} \frac{-2x_1}{x_2^2} \\ \frac{-2x_1}{x_2^2} \frac{2x_1^2}{x_2^3} \end{bmatrix} \tag{44}$$

This function is convex when its hessian is positive semidefinite, which is satisfied when $\frac{2}{x_2} \geq 0$, $\frac{2x_1^2}{x_2^3} \geq 0$ and $\frac{4x_1^2}{x_2^4} - \frac{4x_1^2}{x_2^4} \geq 0$. These are always true and therefore f is convex.

2.8 Exercise 3.18

Not attempted

2.9 Exercise 3.24

Finding if the below probability simplexes are convex, quasiconvex, concave, quasiconcave

2.9.1 quartile

$$\mathbf{quartile}(x) = \inf\{\beta | \mathbf{prob}(x \leq \beta) \geq 0.25\} \tag{45}$$

This function is quasiconvex and quasiconcave. It is not convex since it is not continuous

2.10 Exercise 3.36

Deriving conjugates of the function

2.10.1 Max function

$$f^*(y) = sup\{y^T x - f(x)\} \tag{46}$$

$$f^*(y) = sup\{y^T x - max\{x_i\}\} \tag{47}$$

$$\frac{d}{dx} y^T x - max\{x_i\} = y- \tag{48}$$

3 Homework 3

3.1 Exercise 3.42

For this problem we have to show that the below function is quasiconcave

$$W(x) = \sup\{T||x_1 f_1(t) + \cdots + x_n f_n(t) - f_0(t)| \leq \epsilon \text{ for } 0 \leq t \leq T\} \tag{49}$$

To prove this, we can find if for each ϵ, the function is concave. The definition of quasiconcavity is that all its superlevel sets are concave. In order to evaluate the superlevel sets, we can flip the infimum and supremum

$$W(x) = \inf\{T||x_1 f_1(t) + \cdots + x_n f_n(t) - f_0(t)| \geq \epsilon \text{ for } 0 \leq t \leq T\} \tag{50}$$

This form of the equation now flips the equal sign so we can evaluate the superlevel sets with the threshold being ϵ. The point-wise infimum of a set of concave functions is a concave function. So, as long as the function inside is a concave function, then this holds. The function there is a linear combination of x and therefore it is convex, and therefore quasiconcave.

3.2 Exercise 3.54

Verifying log concavity of the the function

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \tag{51}$$

3.2.1 Part a

Verifying that $f'' f(x) \leq f'(x)^2, x \geq 0$

$$f'(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \tag{52}$$

$$= \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \tag{53}$$

$$f''(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} e^{\frac{-x^2}{2}} \tag{54}$$

$$= -\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} x \tag{55}$$

$$f''(x) f(x) \leq f'(x)^2 \tag{56}$$

$$-\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \leq (\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}})^2 \tag{57}$$

$$-e^{\frac{-x^2}{2}} x \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \leq (e^{\frac{-x^2}{2}})^2 \tag{58}$$

If $x = 0$, the inequality $0 \leq 1$ holds. If $x > 0$, the left hand side of the inequality will always be negative and the right hand side will be greater than 1. Therefore the statement $f''(x)f(x) \leq f'(x)^2, x \geq 0$ holds.

3.2.2 Part b

Verifying that for any t and x, we have $\frac{t^2}{2} \geq \frac{-x^2}{2} + xt$ I am not sure the most rigorous way to do this, so I will evaluate the signs of each side on every combation of zero, positive, and negative.

- $x = 0, t = 0 \rightarrow 0 \geq 0 =>$ evidently true since zero equals zero
- $x > 0$ or $x < 0, t = 0 \rightarrow 0 \geq -\frac{x^2}{2} =>$ evidently true since a negative number is lower than zero
- $x = 0, t > 0$ or $t < 0 \rightarrow \frac{t^2}{2} \geq 0 =>$ evidently true since a positive number is greater than zero
- $x < 0, t > 0 \rightarrow \frac{t^2}{2} \geq -\frac{x^2}{2} - |x|t$ evidently true since a positive number is greater than a negative number
- $x < 0, t < 0 \rightarrow t^2 + x^2 \geq 2|x||t|$ this is the arithmetic-geometric mean inequality

3.2.3 Part c

Part b can be flipped and applying log to both sides we see that $e^{\frac{-t^2}{2}} \leq e^{\frac{x^2}{2} - xt}$

$$e^{\frac{-t^2}{2}} \leq e^{\frac{x^2}{2} - xt} \tag{59}$$

$$\int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \leq \int_{-\infty}^{x} e^{\frac{x^2}{2} - xt} dt \tag{60}$$

$$\int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \leq e^{\frac{x^2}{2}} \int_{-\infty}^{x} e^{-xt} dt \tag{61}$$

3.2.4 Part d

$$\int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \leq -e^{\frac{-x^2}{2}} x \tag{62}$$

3.3 Exercise 3.57

Showing that the function $f(X) = X^{-1}$ is matrix convex on $\mathbb{S}^n_{++}$ If the matrix X is positive definite, then its inverse is also positive definite and therefore matrix convex.

3.4 Exercise 4.4

3.4.1 Part a

Proving that for any $x \in \mathbb{R}^n$ we have $\bar{x} \in \mathbb{F}$ where $\mathbb{F}$ is

$$F = \{x | Q_i x = x, i = 1, ..., k\} \tag{63}$$

Since $\bar{x} = \frac{1}{k} \sum_{i=1}^{k} Q_i x$ and the group $\mathbb{G}$ is closed under products and inverse, any multiplication of the matrices Q_i and Q_j will result in some other Q_s. Therefore, $\frac{1}{k} \sum_{s=1}^{k} Q_s x = \frac{1}{k} \sum_{i=1}^{k} Q_i x$

3.4.2 Part b

Showing that if $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $\mathbb{G}$-invariant, then $f(\bar{x}) \leq f(x)$.

$$\begin{aligned} f(\bar{x}) &= f(\frac{1}{k}\sum_{i=1}^{k} Q_i x) && (64) \\ &\text{Since f is convex} && (65) \\ f(\frac{1}{k}\sum_{i=1}^{k} Q_i x) &\leq \frac{1}{k}\sum_{i=1}^{k} f(Q_i x) && (66) \\ &\text{Since f is G-invariant} && (67) \\ f(\frac{1}{k}\sum_{i=1}^{k} Q_i x) &\leq \frac{1}{k}\sum_{i=1}^{k} f(x) && (68) \\ f(\bar{x}) &\leq f(x) && (69) \end{aligned}$$

3.4.3 Part c

If every function in the optimization problem below is G-invariant, then there exists an equivalent optimization problem

$$\begin{aligned} \text{minimize } & f_0(Q_i x) \\ \text{subject to } & f_i(Q_i x) \leq 0, i = 1, \ldots, m \\ & Q_i x - x = 0 \end{aligned} \tag{70}$$

3.4.4 Part d

Not attempted

3.5 Exercise 4.8

3.5.1 Part a

$$c^T x + \lambda^T (Ax - b) \tag{71}$$

$$\frac{d}{dx} = c + A^T \lambda = 0 \tag{72}$$

$$x^T c + x^T A^T \lambda - \lambda^T b \tag{73}$$

$$x^T (c + A^T \lambda) - \lambda^T b \tag{74}$$

$$p^* = \lambda^T b \iff c = A^T \lambda \tag{75}$$

Otherwise the problem is either infeasible so the the value would be $-\infty$, or the problem is unbounded and therefore ∞

3.5.2 Part b

$$c^T x + \lambda(a^T x - b) \tag{76}$$
$$c + a\lambda = 0 \tag{77}$$
$$c = a\lambda \tag{78}$$

$$p^* = b\lambda \iff c = a\lambda \tag{79}$$

3.5.3 Part c

If c is positive, then the minimization goes to the lower bound l, otherwise, it goes to the upperbound u if c is negative. All feasible values are optimal if $c = 0$

3.5.4 Part d

$$\begin{aligned} \text{minimize } & c^T x \\ \text{subject to } & \mathbf{1}^T x = 1 \\ & x \succeq 0 \end{aligned} \tag{80}$$

In this case, regardless of the sign of c, the optimal value will be the smallest entry c_{min}. This is because if all entries are positive, then the smallest entry will be maxed out by the corresponding $x_i = 1$ and the rest $x_j \neq i = 0$. And if the entries are negative, then the value is still maxed out at $x_i = 1$.

3.5.5 Part e

Not attempted

3.6 Exercise 4.17

$$\begin{aligned} \text{maximize } & \sum_{j=1}^{n} r_j(x_j) \\ \text{subject to } & Ax \preceq c^{max} \\ & x \succeq 0 \end{aligned} \tag{81}$$

Transforming

$$\begin{aligned} \text{maximize } & p^T x_{under} + p_{disc}^T x_{over} \\ \text{subject to } & Ax \preceq c^{max} \\ & x_{under} - q = x_{over} \\ & x_{under}, x_{over} \succeq 0 \end{aligned} \tag{82}$$

with variables x_under and x_over

4 Homework 4

4.1 Exercise 4.11

The task is to formulate different norms as LPs

4.1.1 Part a

Minimizing $\|Ax - b\|_\infty$. This norm is the same as the maximum entry of $Ax - b$

$$\min \|Ax - b\|_\infty = \min \max\{A_i x - b_i, i = 1, \dots, n\} \tag{83}$$

$$\begin{aligned} \min z \\ \text{subject to } z \geq A_i x - b_i, i = 1, \dots, n \end{aligned} \tag{84}$$

4.1.2 Part b

Minimizing$\|Ax - b\|_1$. The l_1 norm is equal to

$$\sum_i |A_i x - b_i| \tag{85}$$

where A_i is the i-th row of the matrix A. We can turn this into

$$\begin{aligned} \text{minimize } \sum_i |z_i| \\ \text{subject to } z_i = A_i x - b_i; i = 1, \dots, m \end{aligned} \tag{86}$$

Removing the absolute value

$$\begin{aligned} \text{minimize } \sum_i z_i \\ \text{subject to } A_i x - b_i \leq z_i; i = 1, \dots, m \\ b_i - A_i x \leq z_i; i = 1, \dots, m \end{aligned} \tag{87}$$

Vector notation

$$\begin{aligned} \text{minimize } \mathbf{1}^T z \\ \text{subject to } Ax - b \preceq z \\ b - Ax \preceq z \end{aligned} \tag{88}$$

4.1.3 Part c

Minimizing $\|Ax - b\|_1$ subject to $\|x\|_\infty \leq 1$

$$\begin{aligned} \text{minimize } \mathbf{1}^T z \\ \text{subject to } Ax - b \preceq z \\ b - Ax \preceq z \\ \|x\|_\infty \leq 1 \end{aligned} \tag{89}$$

Expanding the l_∞ norm

$$
\begin{aligned}
\text{minimize } & \mathbf{1}^T z \\
\text{subject to } & Ax - b \preceq z \\
& b - Ax \preceq z \\
& x \preceq \mathbf{1} \\
& -x \preceq \mathbf{1}
\end{aligned} \tag{90}
$$

4.1.4 Part d

Minimizing $\|x\|_1$ subject to $\|Ax - b\|_\infty \leq 1$

$$
\begin{aligned}
\text{minimize } & \mathbf{1}^T z \\
\text{subject to } & x \preceq z \\
& -x \preceq z \\
& Ax - b \preceq \mathbf{1} \\
& b - Ax \preceq \mathbf{1}
\end{aligned} \tag{91}
$$

4.1.5 Part e

Minimizing $\|Ax - b\|_1 + \|x\|_\infty$

$$
\begin{aligned}
\text{minimize } & \mathbf{1}^T z + s \\
\text{subject to } & x \preceq s \\
& -x \preceq s \\
& Ax - b \preceq z \\
& b - Ax \preceq z
\end{aligned} \tag{92}
$$

4.2 Exercise 4.16

The minimum fuel optimal control problem as an LP. First, starting with what was given in the problem with minimal changes:

$$
\begin{aligned}
\text{minimize } & \sum_{t=0}^{N-1} f(u(t)) \\
\text{subject to } & x(t+1) = Ax(t) + bu(t) \\
& x(N) = x_{des} \\
& x(0) = 0
\end{aligned} \tag{93}
$$

With the fact that

$$
f(u) = \begin{cases} |u| & |u| \leq 1 \\ 2|u| - 1 & |u| > 1 \end{cases} \tag{94}
$$

This is not a linear equation, but we can manipulate it and find that this is equivalent to

$$
f(y) = \begin{cases} y & y \leq 1 \\ 2y - 1 & y > 1 \end{cases} \tag{95}
$$

$$
\text{subject to } -y \leq u \leq y \tag{96}
$$

This function can be linearized by introducing a new variable as the function and using inequalities.

$$\begin{aligned} f(z) &= z \\ \text{subject to } z &\geq y \\ z &\geq 2y - 1 \\ -y \leq u &\leq y \end{aligned} \tag{97}$$

This can be brought to the original problem and formulated as the below LP

$$\begin{aligned} \text{minimize } &\mathbf{1}^T z \\ \text{subject to } x(t+1) &= Ax(t) + bu(t) \\ x(N) &= x_{des} \\ x(0) &= \mathbf{0} \\ z &\geq y \\ z &\geq 2y - 1 \\ -y \leq u(t) &\leq y \end{aligned} \tag{98}$$

4.3 Exercise 4.29

Verifying convexity or quasiconvexity of the problem

$$\begin{aligned} &\text{maximize } \mathbf{prob}(c^T x \geq \alpha) \\ &\text{subject to } Fx \leq g, Ax = b \end{aligned} \tag{99}$$

In this problem, the constraints are both affine so the whole problem is convex as long as the objective function is convex.

4.4 Exercise 4.30

For this problem, the costs are defined as

$$\begin{aligned} c_1 &= \frac{\alpha_1 T r}{w} \\ c_2 &= \alpha_2 r \\ c_3 &= \alpha_3 r w \\ c_T &= \sum_{i=1}^{3} c_i \end{aligned} \tag{100}$$

As stated, this problem can be modeled as

$$\begin{aligned} \text{maximize } &\alpha_4 T r^2 \\ \text{subject to } \frac{\alpha_1 T r}{w} + \alpha_2 r + \alpha_3 r w &\leq C_{max} \\ T_{min} \leq T &\leq T_{max} \\ r_{min} \leq r &\leq r_{max} \\ w_{min} \leq w &\leq w_{max} \\ w &\leq 0.1r \end{aligned} \tag{101}$$

This is equivalent to the below geometric program

$$
\begin{aligned}
\text{minimize } & \frac{1}{\alpha_4}T^{-1}r^{-2} \\
\text{subject to } & \frac{\alpha_1 Trw^{-1} + \alpha_2 r + \alpha_3 rw}{C_{max}} \leq 1 \\
& \frac{1}{T_{max}}T \leq 1 \\
& T_{min}T^{-1} \leq 1 \\
& \frac{1}{w_{max}}w \leq 1 \\
& w_{min}w^{-1} \leq 1 \\
& \frac{1}{r_{max}}r \leq 1 \\
& r_{min}r^{-1} \leq 1 \\
& 10wr^{-1} \leq 1
\end{aligned}
\tag{102}
$$

4.5 Exercise 5.1

Looking at the optimization problem

$$
\begin{aligned}
\text{minimize } & x^2 + 1 \\
\text{subject to } & (x-2)(x-4) \leq 0
\end{aligned}
\tag{103}
$$

4.5.1 Part a

The feasible set is the interval from $[2,4]$. The optimal value is 5 and the optimal solution is $x = 2$

4.5.2 Part b

Solving for the lagrangian

$$
\begin{aligned}
L(\lambda, x) &= x^2 + 1 + \lambda(x-2)(x-4) \\
g(\lambda) &= \inf_x L(\lambda, x) \\
\frac{d}{dx}x^2 + 1 + \lambda(x-2)(x-4) &= 2x + \lambda(2x-6) \\
2x + \lambda 2x &= 6\lambda \\
x &= \frac{3\lambda}{1+\lambda} \\
g(\lambda) &= (\frac{3\lambda}{1+\lambda})^2 + 1 + \lambda((\frac{3\lambda}{1+\lambda}) - 2)(\frac{3\lambda}{1+\lambda} - 4)
\end{aligned}
\tag{104}
$$

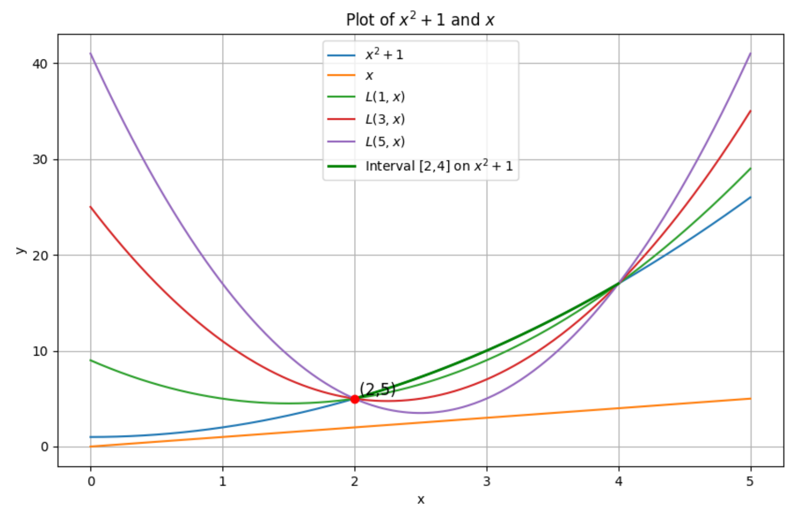


Figure 2: Plot of the feasible set with lagrangian

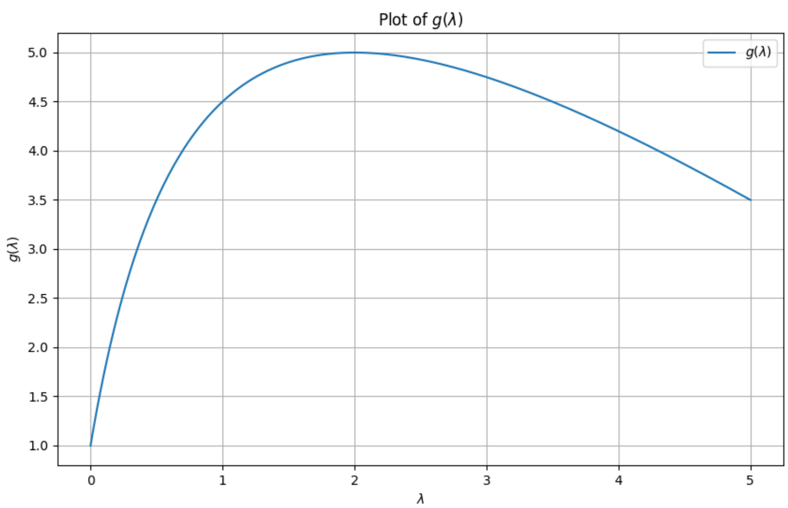


Figure 3: Plot of the lagrangian dual

5 Homework 5

5.1 Exercise 4.15

Relaxation of a boolean LP
A boolean linear program is defined as

$$\begin{aligned} \text{minimize} \quad & c^T x && (105)\\ \text{subject to} \quad & Ax \preceq b && (106)\\ & x_i \in \{0,1\}, i = 1, \dots, n && (107) \end{aligned}$$

This linear program can be relaxed by replacing the integer constraint with linear inequalities

$$\begin{aligned} \text{minimize} \quad & c^T x && (108)\\ \text{subject to} \quad & Ax \preceq b && (109)\\ & 0 \le x_i \le 1, i = 1, \dots, n && (110) \end{aligned}$$

5.1.1 Part a

The optimal value of the LP relaxation is a lower bound on the optimal value. We can use a geometric example to prove this. Since the affine constraints in both the relaxed and original LP form a polyhedron. There are distinct vertex that either do or don't coincide with an integer point. In the event that the polyhedron's vertices land on the integer point, then if the optimal value lies on that vertex, then the optimal value for both the relaxed problem and original problem are equal. In the event that the vertex does not lie on the integer point, then the solution to the relaxed LP is a lower bound on the optimal value of the boolean LP.

The image below illustrates how the vertices lie on integers and therefore the optimal values would be equal. However, if the polyhedron extended out a bit to $(0, 0.5), (0.5, 0)$ as opposed to $(0, 1), (1, 0)$ then the relaxed polyhedron would be a strict lower bound.

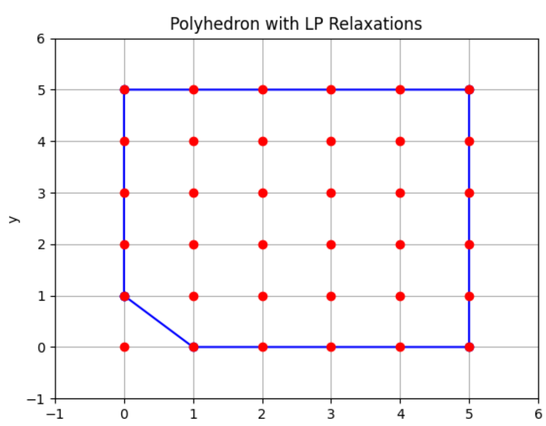


Figure 4: A polyhedron in $\mathbb{R}^2$ with points at each integer

5.1.2 Part b

This section was also answered in part a. However, the solutions for the relaxation and original problem will coincide whenever the polyhedron created by the given A and b lie on an integer.

5.2 Exercise 4.60

- x: An allocation strategy that dictates the portion of the wealth allocated to each i asset. $x \in \mathbb{R}^n$
- n: Number of assets
- N: Number of periods
- $W(t-1)$: Amount of wealth at the beginning of period t or end of period $t-1$
- $\lambda(t)$: Total return during period t. $\lambda(t) = \frac{W(t)}{W(t-1)}$. A random variable with m possible values.
- $\frac{1}{N}\sum_{t=1}^{N} \log \lambda(t)$: Growth rate of the investment over N periods.
- m: Number of deterministic return scenarios.
- π_j: Probability of getting scenario j at any given period. $\pi_j = \mathbf{prob}(\lambda(t) = p_j^T x)$
- p_{ij}: The return for asset i over one period in which scenario j occurs.

Our goal here is to maximize our total expected long-term growth rate. First, for my understanding, I am going to derive the formula for the growth rate.

$$\begin{aligned} W(N) &= W(0) \prod_{t=1}^{N} \lambda(t) \\ \frac{1}{N} \log(W(N) - W(0)) &= \frac{1}{N} \log \prod_{t=1}^{N} \lambda(t) \\ &= \frac{1}{N} \sum_{t=1}^{N} \log \lambda(t) \end{aligned} \tag{111}$$

Now, with this I am going to derive the formula for long term growth rate.

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \log \lambda(t) &= \mathbb{E}[\log \lambda(t)] \\ &= \sum_{j=1}^{m} \pi_j \log p_j^\top x \end{aligned} \tag{112}$$

The final optimization problem is:

$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{m} \pi_j \log p_j^\top x && (113) \\ \text{subject to} \quad & x \succeq 0 && (114) \\ & \mathbf{1}^\top x = 1 && (115) \end{aligned}$$

This question requires that we prove that the above optimization problem is a convex optimization problem. For convention, the problem is turned into

$$
\begin{aligned}
\text{minimize} \quad & -\sum_{j=1}^{m} \pi_j \log p_j^\top x && (116)\\
\text{subject to} \quad & x \succeq 0 && (117)\\
& \mathbf{1}^\top x = 1 && (118)
\end{aligned}
$$

For this to be a convex problem, the objective function and inequalities must be convex, and the equality constraints must be affine.
Firstly, the objective function is convex:

- $p_j^\top x$ is an affine function of the variable x
- $\pi_j \log u$ with $u = (p_j^\top x)$ is a concave function of the variable u, which itself is an affine function of x. The function is then multiplied by a non-negative weight which maintains its concavity
- $\sum v$ with $v = \pi_j \log p_j^\top x$ is a sum of concave functions which is concave.
- $-w$ with $w = \sum_{j=1}^{m} \pi_j \log p_j^\top x$ is the negative of a concave function which is convex.

Clearly, the inequality is convex as it is the cone $\mathbb{R}_+^m$
The equality is indeed affine, therefore the optimization problem is convex!

5.3 Exercise 5.13

5.3.1 Part a

This part involves finding the lagrange dual of the boolean LP, which we can use to get a lower bound on the original optimization problem.

$$
\begin{aligned}
\text{minimize} \quad & c^T x && (119)\\
\text{subject to} \quad & Ax \preceq b && (120)\\
& x_i(1 - x_i) = 0 && (121)
\end{aligned}
$$

$$
\begin{aligned}
L(x, \lambda, \nu) &= c^T x + \lambda^T (Ax - b) + \nu^T (x - \mathbf{diag}(x)x)\\
&= c^T x + \lambda^T Ax - \lambda^T b + \nu^T x - x^T \mathbf{diag}(\nu)x
\end{aligned} \tag{122}
$$

$$
\begin{aligned}
\nabla L(x, \lambda, \nu) &= c + A^T\lambda + \nu - 2\mathbf{diag}(\nu)x\\
2\mathbf{diag}(\nu)x &= c + A^T\lambda + \nu\\
x &= \frac{1}{2}\mathbf{diag}(\nu)^{-1}(c + A^T\lambda + \nu)
\end{aligned} \tag{123}
$$

$$
\begin{aligned}
g(\lambda, \nu) &= -\lambda^T b + (c + A^T\lambda + \nu)^T x - \frac{1}{2}x^T(c + A^T\lambda + \nu)\\
&= -\lambda^T b + \frac{1}{2}(c + A^T\lambda + \nu)^T \frac{1}{2}\mathbf{diag}(\nu)^{-1}(c + A^T\lambda + \nu)\\
&= -\lambda^T b + \frac{1}{4}(c + A^T\lambda + \nu)^T (\mathbf{diag}(\nu)^{-1}(c + A^T\lambda + \nu))
\end{aligned} \tag{124}
$$

$$g(\lambda, \nu) = \begin{cases} -\lambda^T b + \frac{1}{4} z^T \tilde{D} z & \nu \succeq 0 \\ -\infty & o/w \end{cases} \tag{125}$$

where $z = (c + A^T \lambda + \nu)$ and $\tilde{D} = \mathbf{diag}(\nu)^{-1}$
Solving for the maximum here we have the optimization problem

$$\begin{aligned} \text{maximize} \quad & g(\lambda, \nu) = -\lambda^T b - \frac{1}{4}(c + A^T \lambda + \nu)^T (\mathbf{diag}(\nu)^{-1}(c + A^T \lambda + \nu)) && (126) \\ \text{subject to} \quad & \lambda \succeq 0 && (127) \\ & \nu \succeq 0 && (128) \end{aligned}$$

5.3.2 Part b

Now, first we derive the lagrangian dual of the LP relaxation

$$\begin{aligned} \text{minimize} \quad & c^T x && (129) \\ \text{subject to} \quad & Ax \preceq b && (130) \\ & \mathbf{0} \preceq x \preceq \mathbf{1} && (131) \end{aligned}$$

$$\begin{aligned} L(x, \lambda, \nu_1, \nu_2) &= c^T x + \lambda^T (Ax - b) - \nu_1^T (x) + \nu_2^T (x - 1) \\ &= (c + A^T \lambda - \nu_1 + \nu_2)^T x - \lambda^T b - \mathbf{1}^T \nu_2 \end{aligned} \tag{132}$$

$$\begin{aligned} \nabla L(x, \lambda, \nu_1, \nu_2) &= c + A^T \lambda - \nu_1 + \nu_2 \\ 0 &= c + A^T \lambda - \nu_1 + \nu_2 \end{aligned} \tag{133}$$

$$g(\lambda, \nu_1, \nu_2) = \begin{cases} -\lambda^T b - \mathbf{1}^T \nu_2 & 0 = c + A^T \lambda - \nu_1 + \nu_2 \\ -\infty & o.w. \end{cases} \tag{134}$$

The optimization problem with this lagrangian is

$$\begin{aligned} \text{maximize} \quad & g(\lambda, \nu_1, \nu_2) = -\lambda^T b - \mathbf{1}^T \nu_2 && (135) \\ \text{subject to} \quad & 0 = c + A^T \lambda - \nu_1 + \nu_2 && (136) \\ & \lambda \succeq 0, \nu_1 \succeq 0, \nu_2 \succeq 0 && (137) \end{aligned}$$

That is apparently the same as the above but I can't quite understand why

5.4 Exercise 6.2

This exercise involves finding the solution of the scalar norm approximation problem with l_1, l_2, l_∞

$$\text{minimize} \quad \|x\mathbf{1} - b\| \tag{138}$$

Where $x \in \mathbb{R}$ and $b \in \mathbb{R}^n$

5.4.1 Part a

l_2 norm

$$\text{minimize} \quad \|x\mathbf{1} - b\|_2 \tag{139}$$

$$\begin{aligned} f(x) &= \frac{1}{n}\sum_{i=1}^{n}(x-b_i)^2 \\ \nabla f(x) &= \frac{2}{n}\sum_{i=1}^{n}(x-b_i) \\ \sum_{i=1}^{n} b_i &= nx \\ x &= \frac{1}{n}\sum_{i=1}^{n} b_i \text{ or } \frac{\mathbf{1}^T b}{n} \end{aligned} \tag{140}$$

5.4.2 Part b

l_1 norm

$$\text{minimize} \quad \|x\mathbf{1} - b\|_1 \tag{141}$$

$$f(x) = \frac{1}{n}\sum_{i=1}^{n}|x-b_i| \tag{142}$$

The answer to this is the median of the vector b.

5.4.3 Part c

l_1 norm

$$\text{minimize} \quad \|x\mathbf{1} - b\|_\infty \tag{143}$$

Here we are minimizing the maximum absolute distance between any of the components of b and x

$$\max_i\{|x-b_i|\} \tag{144}$$

The only two important numbers in b are b_{min}, b_{max}. In order to minimize the distance between x and these two points, x is set to the midrange between those two points

$$x = \frac{b_{min} + b_{max}}{2} \tag{145}$$

5.5 Additional Problem HW5.2

Formulating different optimization problems as semidefinite programs with the variable $x \in \mathbb{R}^n$ and

$$F(x) = F_0 + x_1F_1 + x_2F_2 + \dots x_nF_n \tag{146}$$

where $F_i \in \mathbb{S}^m$.

A semidefinite program is an optimization problem in the form

$$\begin{aligned} \text{minimize} \quad & c^T x && (147)\\ \text{subject to} \quad & x_1 F_1 + x_2 F_2 + \dots x_n F_n + G \preceq 0 && (148)\\ & Ax = b && (149) \end{aligned}$$

where $F_i, G \in \mathbb{S}^k$

5.5.1 Part a

Minimize $f(x) = c^T F(x)^{-1} c$ where $c \in \mathbb{R}^m$:

$$\begin{aligned} \text{minimize} \quad & y && (150)\\ \text{subject to} \quad & y \geq c^T F(x)^{-1} c && (151) \end{aligned}$$

$$\begin{aligned} y &\geq c^T F(x)^{-1} c \\ yF(x) &\succeq cc^T \\ yF(x) - cc^t &\succeq 0 \end{aligned} \tag{152}$$

$$\begin{bmatrix} F(x) & c \\ c^T & y \end{bmatrix} \succeq 0 \tag{153}$$

5.5.2 Part b

Minimize $f(x) = \max_i c_i^T F(x)^{-1} c_i$

$$\begin{aligned} \text{minimize} \quad & y && (154)\\ \text{subject to} \quad & y \geq c_i^T F(x)^{-1} c_i \forall i \in \{1, \dots, K\} && (155) \end{aligned}$$

$$\begin{bmatrix} F(x) & c_i \\ c_i^T & y \end{bmatrix} \succeq 0 \forall i \tag{156}$$

5.5.3 Part c

Minimize $f(x) = \sup_{\|c\|_2 \leq 1} c^T F(x)^{-1} c$

$$\begin{aligned} \text{minimize} \quad & \text{maximize } y && (157)\\ \text{subject to} \quad & y \geq c^T F(x)^{-1} c && (158)\\ & \|c\|_2 \leq 1 && (159) \end{aligned}$$

Don't really get this one but the answer is

$$\begin{bmatrix} F(x) & I \\ I & yI \end{bmatrix} \succeq 0 \tag{160}$$

5.6 Additional Problem HW5.5

Maximizing house profit in a gamble and imputed probabilities.

- n: number of participants
- m: number of outcomes that the participants can bet on
- q_i: number of gambling contracts that a participant can bet on $q_i > 0$
- p_i: price of gambling contract i
- $\mathbb{S}_i$: $\mathbb{S}_i \subset \{1, \ldots, m\}$ A set the participant i specifies that they are betting the outcome will lie in.
- x_i: The number of contracts that the house sells to the participant. If the true outcome j is in $\mathbb{S}_i$, the partipant will receive \$1 per x_i the house has sold them. $0 \leq x_i \leq q_i$
- $R(x)$: Revenue that the house collects. $R(x) = p^T x$ or $R(x) = p_1 x_1 + \cdots + p_n x_n$
- $C(x)$: Cost that the house has to pay out to the participants that depends on outcome j. $C(x) = \sum_{j \in \mathbb{S}_i} x_i$ which is a random variable that depends on the outcome.
- $P(x)$: House profit. $R(x) - C(x)$

5.6.1 Part a

This section involves finding the optimal worst-case house strategy.
Worst case profit is defined as below

$$P_{wc}(x) = p^T x - \max_j \sum_{j \in \mathbb{S}_i} x_i \tag{161}$$

The house wants to solve the following optimization problem.

$$\begin{aligned} \text{maximize} \quad & p^T x - \max_j \sum_{j \in \mathbb{S}_i} x_i && (162) \\ \text{subject to} \quad & 0 \preceq x \preceq q && (163) \end{aligned}$$

We can create a series of vectors $z_i \in \mathbb{R}^m$ where each entry is 1 if the outcome j is in the set $\mathbb{S}_i$ and it is 0 otherwise. The multiplication of the two would correspond to a vector with each outcome stored in a vector. $z* = \max_j \{x_i z_i\} \forall i$. This represents the maximum loss on each participant. The optimization problem is then turned into

$$\begin{aligned} \text{minimize} \quad & p^T x - \mathbf{1}^T z^* && (164) \\ \text{subject to} \quad & 0 \preceq x \preceq q && (165) \\ & z_i^* \geq \max_j \{x_i \mathbf{z}_i\} \forall i && (166) \end{aligned}$$

With crazy confusing notation.

5.6.2 Part b

I am going to borrow a little from the solutions notation to turn mine into an LP.
We can turn the answer from part a into an LP by collecting the previous vector into a matrix Z and introducing a new scalar variable.

$$\begin{aligned} \text{minimize} \quad & p^T x - z && (167)\\ \text{subject to} \quad & z\mathbf{1} \succeq Zx && (168)\\ & 0 \preceq x \preceq q && (169) \end{aligned}$$

We can now find the dual from this LP

$$\begin{aligned} L(x, z, \lambda_1, \lambda_2, \lambda_3) &= p^T x - z + \lambda_1^T (Zx - z\mathbf{1}) - \lambda_2^T (x) + \lambda_3^T (x - q) \\ &= (p + Z^T \lambda_1 - \lambda_2 + \lambda_3)^T x - z + z\lambda_1^T \mathbf{1} - \lambda_3^T q \end{aligned} \tag{170}$$

$$\begin{aligned} \nabla_x L(x, z, \lambda_1, \lambda_2, \lambda_3) &= p + Z^T \lambda_1 - \lambda_2 + \lambda_3 \\ 0 &= p + Z^T \lambda_1 - \lambda_2 + \lambda_3 \\ \nabla_z L(x, z, \lambda_1, \lambda_2, \lambda_3) &= -1 + \lambda_1^T \mathbf{1} \\ 1 &= \lambda_1^T \mathbf{1} \end{aligned} \tag{171}$$

$$g(\lambda_1, \lambda_2, \lambda_3) = \begin{cases} -\lambda_3^T q & 1 = \lambda_1^T \mathbf{1}, 0 = p + Z^T \lambda_1 - \lambda_2 + \lambda_3 \\ 0 & o.w. \end{cases} \tag{172}$$

The dual optimization problem is as follows

$$\begin{aligned} \text{maximize} \quad & -\lambda_3^T q && (173)\\ \text{subject to} \quad & 0 = p + Z^T \lambda_1 - \lambda_2 + \lambda_3 && (174)\\ & 1 = \lambda_1^T \mathbf{1} && (175)\\ & \lambda_1 \succeq 0, \lambda_2 \succeq 0, \lambda_3 \succeq 0 && (176) \end{aligned}$$

$\lambda_1 \in \mathbb{R}^n$ is a probability distribution as it sums up to 1 and each value is non-negative. Therefore, we set $\pi = \lambda_1^*$ to get the probability distribution that maximizes both expected and worst case scenario.