.

CME307 HW1

Jacob Perricone
Stanford University
jacobp2@stanford.edu

1 Problem 1

1. Consider the feasible set $F := \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ Where data matrix $A^{m \times n}$ and $b \in \mathbb{R}^m$. Prove that F is a convex set.

To prove that F is convex we must show that the point $z = \alpha x_1 + (1-\alpha)x_2 \in F$ with $0 \leq \alpha \leq 1$ and $x_1, x_2 \in F$. Now since both $x_1, x_2 \in F$ and $\alpha \geq 0$, it follows that $x_1, x_2 \geq 0$ and thus $z \geq 0$. Now to prove $Az = b$, notice:

$$Az = A[\alpha x_1 + (1-\alpha)x_2]$$
$$Az = \alpha Ax_1 + (1-\alpha)Ax_2 = \alpha b + (1-\alpha)b = b$$

Thus $z \in F$ and F is convex

---

2. Fix data matrix A and consider the b-data set for F defined above:

$$B := \{b \in \mathbb{R}^m : \text{F is not empty}\}$$

Choose $b_1, b_2 \in B$ s.t $b_1 \neq b_2$ and let x_1, x_2 be the corresponding points in F with $Ax_i = b_i, i \in \{1, 2\}$. Now define $c = \alpha b_1 + (1-\alpha)b_2$ where $0 \leq \alpha \leq 1$. Now for B to be convex, $c \in B$, or, equivalently, $\exists z \in F$ s.t $Az = c, z \geq 0$:

$$\begin{aligned} c &= \alpha b_1 + (1-\alpha)b_2 \\ &= \alpha Ax_1 + (1-\alpha)Ax_2 \\ &= A(\alpha x_1 + (1-\alpha)x_2) \\ z &= \alpha x_1 + (1-\alpha)x_2 \in F \text{ by the convexity of F} \\ c &= Az, z \in F \end{aligned}$$

Thus $c \in B$ and B is convex.

---

3. Fix data matrix A and consider the linearly constrained convex minimization problem

$$z(b) := \min f(x)$$
$$\text{s.t } Ax = b, x \geq 0$$

where $f(x)$ is a convex function and the minimal value function $z(b)$ is an implicit function of b. Prove that $z(b)$ is a convex function of $b \in B$.

Choose b_1, b_2 be 2 solutions to the minimization problem above, and let x_1^*, x_2^* be defined as

$$\min_x f(x)$$
$$\text{s.t } Ax = b_i, x \geq 0$$
$$i \in \{1, 2\}$$

i.e. the x values where $f(x)$ attains a minimum for the optimization problem bounded by b_i. Define:

$$c = tb_1 + (1-t)b_2$$

where $0 \leq t \leq 1$ and $z = tx_1^* + (1-t)x_2^*$. Since $f(x)$ is a convex function,

$$f(z) \leq tf(x_1^*) + (1-t)f(x_2^*)$$

Now since $z(b_i) = f(x_i^*)$ by definition and $z(c) \leq f(x) \,\forall\, x : \{Ax = c, x \geq 0\}$ by definition, convexity follows by:

$$f(z) \leq tf(x_1^*) + (1-t)f(x_2^*)$$
$$z(c) \leq f(z) \leq tz(b_1) + (1-t)z(b_2)$$

Problem 2

Show that the the dual cone of the n-dimensional nonnegative orthant cone $\mathbb{R}_+^n$ is itself, that is

$$(\mathbb{R}_+^n)^* = \mathbb{R}_+^n$$

It is clear that the nonnegative orthant $\mathbb{R}_+^n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq 0\}$ is a closed convex cone. The dual orthant is defined as: $(\mathbb{R}_+^n)^* := \{\mathbf{z}^T\mathbf{x} \geq 0, \forall \mathbf{x} \geq 0\}$.This set clearly contains the nonnegative orthant: $\{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} \geq 0\}$. Thus $R_+^n \subset (\mathbb{R}_+^n)^*$. It is also clear that the dual of the nonnegative orthant contains the nonnegative orthant and **nothing more**. To see this let $\mathbf{z} \in \mathbb{R}^n$, with $z_i < 0$, for some i. Notice that $\mathbf{z}^T\mathbf{e}_i < 0$, where e_i is the $i-th$ unit vector (which is a member of $(\mathbb{R}_+^n)$). Thus z is not a member of the dual cone and no $\mathbf{z}$ with a $z_i < 0$ is in $(\mathbb{R}_+^n)^*$, thus $(\mathbb{R}_+^n)^* \subset \mathbb{R}_+^n$.

Problem 3

Using Theorem 5 in Lecture Note 1 to prove that the linear system

$$A^TAx = A^Tb$$

always has a solution x for any given matrix $A \in \mathbb{R}^{m\times n}$ and vector $b \in \mathbb{R}^m$.

To show that the system above always has a solution, using Theorem 5 in the Notes we must show that the system

$$A^TAy = 0 \quad b^TAy \neq 0$$

always does not have a solution. Now let us assume that $A^TAy = 0$, meaning that Ay is in the span of A and as well orthogonal to A. Thus it must hold that $Ay = 0$. Now, we must have that $b^TAy \neq 0$, but $b^TAy = b^T\hat{0} = 0$, Thus the system above never has a solution.

Problem 4

Let $g_1, \dots, g_m$ be a collection of concave functions on $\mathbb{R}^n$ such that

$$S = \{x : g_i(x) > 0, \text{ for } i = 1, \dots, m\} \neq 0$$

Show that for any positive constant μ and any convex function f on $\mathbb{R}^n$ the function

$$h(x) = f(x) - \mu \sum_{i=1}^{m} \log[g_i(x)]$$

is convex over S.

Take two points x_1, x_2 in S so that

$$h(x_1) = f(x_1) - \mu \sum_i \log[g_i(x_1)]$$
$$h(x_2) = f(x_2) - \mu \sum_i \log[g_i(x_2)]$$

To prove convexity we have to show that for $0 \leq t \leq 1$, $h(tx_1 + (1-t)x_2) \leq th(x_1) + (1-t)h(x_2)$. Now

$$h(tx_1 + (1-t)x_1) = f(tx_1 + (1-t)x_2) - \mu \sum_i \log[g_i(tx_1 + (1-t)x_2)]$$

By the convexity of f we know that

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

Thus

$$h(tx_1 + (1-t)x_1) \leq tf(x_1) + (1-t)f(x_2) - \mu \sum_i \log[g_i(tx_1 + (1-t)x_2)]$$

Moving onto the the functions g_i. Now since g_i is concave and log is monotonically increasing

$$g_i(x_1 t + (1-t)x_2) \geq g_i(x_1)t + (1-t)g_i(x_2)$$
$$\log[g_i(x_1 t + (1-t)x_2)] \geq \log[g_i(x_1)t + (1-t)g_i(x_2)]$$

Now by the concavity of the log function

$$\begin{aligned} \log[g_i(x_1 t + (1-t)x_2)] &\geq \log[g_i(x_1)t + (1-t)g_i(x_2)] \\ &\geq t\log[g_i(x_1)] + (1-t)\log[g_i(x_2)] \end{aligned}$$

Since the inequality applies for each term within the sum and $\mu > 0$, one concludes that

$$-\mu \sum_i \log[g_i(tx_1 + (1-t)x_2)] \leq -\mu \left[t \sum_i \log[g_i(x_1)] + (1-t) \sum_i \log[g_i(x_2)] \right]$$

Putting them all together we have that

$$\begin{aligned} h(tx_1 + (1-t)x_1) &\leq t\big[f(x_1) - \mu \sum_i \log[g_i(x_1)]\big] + (1-t)\big[f(x_2) - \mu \sum_i \log[g_i(x_2)]\big] \\ &\leq th(x_1) + (1-t)h(x_2) \end{aligned}$$

and the proof is complete.

Problem 5

Consider the min-risk portfolio management problem in Lecture Note 2.

$$\min x^T V x$$
$$\text{s. t} r^T x \geq \mu,$$
$$e^T x = 1, x \geq 0$$

where data vector $r \in \mathbb{R}^n$ representing expected return of n stocks, and $V \in \mathbb{R}^{n \times n}$. representing co-variance matrix of n stocks, μ representing the desired return of an investment portfolio, and e is the vector of all ones. The decision problem is to allocate a total 100% of asset to each stock to minimize the risk while keep the desired return. Thus $x_i, i = 1 \ldots, n$ represents the percentage of the total asset invested in stock i.

1. Now, suppose for simplicity the company's policy is to invest in each stock at one of the three levels:$.05, .1, .2$.. How to add constraints to enforce this policy.

Define 3 binary variables y_{i1}, y_{i2}, y_{i3} and let $x_i = .05y_{i1} + .1y_{i2} + .2y_{i3}$ and $\sum_j y_{ij} \leq 1$. To enforce binary variables, add the constraint that $y_{ij}^2 - y_{ij} = 0$. Thus the added constraints are

$$x_i = .05y_{i1} + .1y_{i2} + .2y_{i3}$$
$$\sum_j y_{ij} \leq 1 \text{ for } i \in 1, \ldots, n$$
$$y_{ij}^2 - y_{ij} = 0 \text{ for } i \in 1, \ldots, n \text{ and } j \in 1, 2, 3$$

---

2. Suppose that the company also does not want to invest in more than 20 stocks. How to add constraints to enforce this additional policy?

Simply add the constraint below to the ones above:

$$\sum_i \sum_j y_{ij} \leq 20$$

Problem 6

Consider the SOCP problem described in in Lecture Note 3:

$$\min 2x_1 + x_2 + x_3$$
$$\text{s.t } x_1 + x_2 + x_3 = 1$$
$$x_1 - \sqrt{x_2^2 + x_3^2} \geq 0$$

1. Show that the feasible region is a convex set.

It suffices to show that each constraint is a convex set. Then, we know that the intersection of any two convex sets is convex. The first constraint $x_1 + x_2 + x_3 = 1$ is a hyperplane in $\mathbb{R}^3$. It is well known that hyperplanes are convex, thus, for future reference, I will prove this in $\mathbb{R}^n$. A hyperplane $\mathbb{R}^n$ is defined by

$H := \{x : a^T x = c, a, x \in \mathbb{R}^n\}$. Now choose $x^1, x^2 \in H$. and define $z = (1-t)x^1 + tx^2, 0 \leq t \leq 1$. We must prove that $z \in H$. Note that:

$$\begin{aligned} z &= \sum_{i=1}^{n} a_i (1-t) x_i^1 + t x_2^i \\ &= \sum_{i=1}^{n} a_i x_1^i (1-t) + \sum_{i=1}^{n} a_i x_2^i t \\ &= (1-t) a^T x_1 + t a^T x_2 = (1-t)c + tc = c \in H \end{aligned}$$

Now the second constraint is an ice-cream cone in $\mathbb{R}^3$. Generally an icecream cone is define as the set

$$C = \{(x, r) \in \mathbb{R}^{n-1} \times \mathbb{R} : ||x||_2 \leq r\}$$

To prove C is convex, choose $x_1, x_2 \in C$ and as always let $z = tx_1 + (1-t)x_2, t \in [0, 1]$. One can see using the triangle inequality

$$\begin{aligned} ||z|| &= ||tx_1 + (1-t)x_2||_2 \\ &\leq t||x_1||_2 + (1-t)||x_2||_2 \\ &\leq tr + (1-t)r \\ &= r \end{aligned}$$

thus $z \in C$, so C is convex and the intersection of the two is convex.

---

2. Try to find a minimizer of the problem and argue why it is a minimizer.

$$\begin{aligned} \min\ & 2x_1 + x_2 + x_3 \\ \text{s.t } & x_1 + x_2 + x_3 = 1 \\ & x_1 - \sqrt{x_2^2 + x_3^2} \quad \geq 0 \end{aligned}$$

Now notice that the optimization problem above can be rewritten as

$$\begin{aligned} \min 1 + x_1 \\ x_1 + x_2 + x_3 = 1 \\ x_1 \geq \sqrt{x_2^2 + x_3^2} \end{aligned}$$

Now we want to minimize x_1, which can be accomplished by minimizing $x_1 = \sqrt{x_2^2 + x_3^2}$. Now thinking about this as a triangle we have that

$$\begin{aligned} x_2 = x_1 \cos\theta \\ x_3 = x_1 \sin\theta \end{aligned}$$

Thus the constraint can be reformulated so that

$$x_1 + x_2 + x_3 = 1 = x_1 + x_1 \cos\theta + x_1 \sin\theta$$

Taking the derivatives θ we have

$$0 = x_1' + x_1' \cos\theta - x_1 \sin\theta + x_1 \sin\theta + x_1 \cos\theta$$

Now, setting $x_1' = 0$ and noting that $x_1 \neq 0$. Thus

$$\begin{aligned} -x_1 \sin\theta + x_1 \cos\theta = 0 \\ x_1(\cos\theta - \sin\theta) = 0 \end{aligned}$$

Thus $\sin\theta = \cos\theta \to \theta = \frac{\pi}{4}$.

Now plugging in we have:

$$x_1 + x_2 + x_3 = 1$$
$$x_1 + x_1 \cos\frac{\pi}{4} + x_1 \sin\frac{\pi}{4} = 1$$
$$x_1(1+\sqrt{2}) = 1$$
$$x_1 = \sqrt{2} - 1$$

Thus $x_2 = x_3 = 1 - \frac{\sqrt{2}}{2}$

Problem 7

Prove that the set $C := \{Ax : x \geq 0 \in \mathbb{R}^n\}$ is closed and convex cone.

It is pretty obvious that the set C is convex since taking $x_1, x_2 \in C$, we have that $z = tx_1 + (1-t)x_2, t \in [0,1] \in C$. Now to prove that it is closed, recall Caratheodorys theorem, which states that a polyhedral cone can be generated by a set of basic directional vectors. Specifically, Caratheodory's theorem states that for $A \in \mathbb{R}^{m\times n}$ where $n > m$, any b in the the polyhedral cone $C = \{Ax : x \geq 0\}$ can be written as a linear combination of linearly independent vectors in the vector space of A. Formally, for any $b \in C$

$$b = \sum_{j=1}^{d} a_{j_i} x_{j_i}, x_{j_i} \geq 0 \forall i$$

for some linearly independent vectors $a_{j_1}, \ldots, a_{j_d}$ chosen from $a_1, \ldots, a_n$.

More generally, the proof can be rephrased as, for any $a_1 \ldots a_n$ vectors in a real vector space V, if x is a linear combination of the a_i with non-negative coefficients, then x is a linear combination with non-negative coefficients of a linearly independent subset of the a_i. The cone generated by the vectors $a_1, \ldots, a_n$, is defined as all positive linear combinations of the vectors $a_i, i = 1, \ldots, n$. I would like to derive this theorem for educational purposes, so I am going to go through a proof, even though the notes assume we can use them. First let $S = \{a_1, \ldots, a_n\}$ be a finite subset of a vector space X. The convex cone C generated by S is given by

$$C = \{\sum_{i=1}^{m} \lambda_i a_i, \lambda_i \geq 0 \forall i\}$$

In other words, the convex cone, C, generated by S is the smallest convex cone that includes S. Now set a $x \in C$ so that $x = \sum_{i=1}^{k} \lambda_i a_i$. We must prove that there is a linearly independent subset L of S such that for non-negative constants $\alpha_l : l \in L$ so that $x = \sum_{l \in L} \alpha_l l$. Now only $\lambda_i > 0$ matter (we can just drop $\lambda_i = 0$ terms). If S is linearly independent, we are done. Otherwise, we know that there exists scalars $\theta_1, \ldots, \theta_n$, that are all not zero, such that $\sum_{i=1}^{n} \theta_i a_i = 0$. If all θ_i are negative, just multiply by -1. Now let $\psi = \max\{\frac{\theta_i}{\lambda_i} : i = 1, \ldots, n\}$. Thus we have $\psi > 0, \lambda_i \geq \frac{\theta_i}{\psi}$ for all i and that for some $i : \lambda_i = \frac{\theta_i}{\psi}$. Then x can be represented as

$$x = \sum_{i=1}^{n} \lambda_i a_i = \sum_{i=1}^{n} (\lambda_i - \frac{\theta_i}{\psi}) a_i$$

Implying that x can be rewritten as a linear combination of non-negative coefficients with one of them being zero. Thus if S is not linearly independent, one can write x as a linear combination of positive coefficients of $n-1$ vectors of S. We can repeat this process until all the vectors are linearly independent. This completes the proof of caratheodory's theorem. Now, assume that a sequence $\{y_n\}$ in K satisfies that $y_n \to y \in X$, that is it converges in our vector space. Since the collection of linearly independent subsets of S is a finite set, by the discussion above we can find a linearly independent subset of S, $\{v_1, \ldots, v_k\} \in S, k \leq n$, and a subsequence $\{y_n\}$ such that

$$y_n = \sum_{i=1}^{k} \lambda_i^n v_i$$

with all λ_i^n non-negative. Now it follows that since the linear span of $\{v_1, \ldots, v_k\} \in S$, call it Y, is a closed vector subspace of X, then there exists $\lambda_1, \ldots, \lambda_k, \lambda_i \geq 0$ such that $y = \sum_{i=1}^{K} \lambda_i v_i$. Now for each $y \in Y$, let $||y|| = \sum_{i=1}^{m} |\lambda_i|$. Now since Y is closed:

$$||y_n - y|| = || \sum_{i=1}^{k} \lambda_i^n v_i - \sum_{i=1}^{k} \lambda_i v_i = \sum_{i=1}^{k} |\lambda_i^n - \lambda| \to 0$$

so that $\lambda_i^n \to \lambda_i$ for all i. Hence $y \in K$ and thus K is closed.

Problem 8

1. Prove Gordon's Lemma that exactly one of the two systems has a solution

$$Ax > 0$$
$$y^T A = 0, y \geq 0, y \neq 0$$

Let $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b, y \in \mathbb{R}^m$. We can then rewrite $Ax > 0$

$$\begin{aligned} 0 &= A(x' - x'') - s, \quad s, x', x'' \geq 0 \\ &= Ax' - Ax'' - s \\ &= [A \quad -A \quad -I] \begin{bmatrix} x' \\ x'' \\ s \end{bmatrix} \end{aligned}$$

Now let $A' = [A \quad -A \quad -I]$ and $z' = \begin{bmatrix} x' \\ x'' \\ s \end{bmatrix}$. The system above then becomes

$$A'z' = b, z' \geq 0$$

which is identically to the system in farkas lemma. Now we must show $A^T y = 0, y \neq 0, y \geq 0$, is equivalent to the alternative system $(-A'^T y \geq 0, b^T y > 0)$ given farkas lemma. Substituting in A' yields

$$\begin{aligned} \begin{bmatrix} -A^T & A^T & I \end{bmatrix} y' &\geq 0 \\ -A^T y' &\geq 0 \\ A^T y' &\geq 0 \\ y' &\geq 0 \\ b^T y' &> 0 \end{aligned}$$

Since both $-A^T y' \geq 0$ and $A^T y' \geq 0$ it follows that $A^T y' = 0$. Also since $y' \geq 0$, and $b^T y' > 0$, it follows that $y' \neq 0$. Thus setting $y' = y$, the alternative system above can be rewritten as

$$A^T y = 0, y \geq 0, y \neq 0$$

Thus we have two alternative system pairs and by Farkas lemma only one can have a solution

---

2. Stiemke's theorem: Prove exactly one of the two below has a solution:

$$\begin{aligned} i) &\quad Ax \geq 0, Ax \neq 0 \\ ii) &\quad A^T y = 0, y > 0 \end{aligned}$$

Let $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b, y \in \mathbb{R}^m$. Assume that $Ax \geq 0$ and $Ax \neq 0$. Thus $Ax > 0$. Following from above we have in part 1. We can rewrite $Ax > 0$ as:

$$\begin{aligned} 0 &= A(x' - x'') - s, \quad s, x', x'' \geq 0 \\ &= Ax' - Ax'' - s \\ &= \begin{bmatrix} A & -A & -I \end{bmatrix} \begin{bmatrix} x' \\ x'' \\ s \end{bmatrix} \end{aligned}$$

Now let $A' = [A \quad -A \quad -I]$ and $z' = \begin{bmatrix} x' \\ x'' \\ s \end{bmatrix}$. The system above then becomes

$$A'z' = b, z' \geq 0$$

which is identically to the system in farkas lemma.
Now we must show $A^T y = 0, y > 0$, is equivalent to the alternative system $(-A'^T y \geq 0, b^T y > 0)$ given farkas lemma. Substituting in A' yields

$$\begin{aligned} \begin{bmatrix} -A^T & A^T & I \end{bmatrix} y' &\geq 0 \\ -A^T y' &\geq 0 \\ A^T y' &\geq 0 \\ y' &\geq 0 \\ b^T y' &> 0 \end{aligned}$$

Since both $-A^T y' \geq 0$ and $A^T y' \geq 0$ it follows that $A^T y' = 0$. Also since $y' \geq 0$, and $b^T y' > 0$, it follows that $y' \neq 0$, so we can just constrain $y' > 0$. Thus setting $y' = y$, the alternative system above can be rewritten as

$$A^T y = 0, y > 0$$

Thus we have two alternative system pairs and by Farkas lemma only one can have a solution

---

3. Gale's Theorem: Prove only one of the following has a solution

$$\begin{aligned} i) &\quad Ax \leq b \\ ii) &\quad A^T y = 0, y^T b < 0, y \geq 0 \end{aligned}$$

Rewriting the inequality $Ax \leq b$ as:

$$\begin{aligned} &A(x' - x'') + s = bx', x'', s \geq 0 \\ &Ax' - Ax'' + s = b \end{aligned}$$

Now suppose that $A = B^T$, where $B \in \mathbb{R}^{n \times m}$, we can then rewrite the system above as

$$\begin{bmatrix} B^T & -B^T & I \end{bmatrix} \begin{bmatrix} x' \\ x'' \\ s \end{bmatrix} = b$$

Let $A' = \begin{bmatrix} B^T & -B^T & I \end{bmatrix}$ and $z' = \begin{bmatrix} x' \\ x'' \\ s \end{bmatrix}$. The system can then be reformulated as:

$$A'z' = b, z' \geq 0$$

which is identical to one of the systems in Farkas lemma. Now we must $A^T y = 0, y^T b < 0, y \geq 0$, can be written using the alternative system $(-A'^T y \geq 0, b^T y > 0)$ in farkas lemma. Substituting in A' yields

$$\begin{aligned} \begin{bmatrix} -B^T & B^T & -I \end{bmatrix} y' &\geq 0 \\ -By' &\geq 0 \\ By' &\geq 0 \\ -y' &\geq 0 \\ b^T y' &> 0 \end{aligned}$$

Since both $-By' \geq 0$ and $B^T y' \geq 0$ it follows that $By' = 0$. Now suppose that $-y' = y$, the alternative system above can be rewritten as

$$\begin{aligned} By &= 0 \\ y &\leq 0 \\ b^T y &< 0 \end{aligned}$$

Substituting $A = B^T$ gives:

$$\begin{aligned} A^T y &= 0 \\ y &\geq 0 \\ b^T y &< 0 \end{aligned}$$

Thus we have two alternative system pairs and by Farkas lemma only one can have a solution

Problem 9

See group type up submitted by Weronicka.