MGMT 690 - Pset 1

Spring 2024

Instructions:

- This pset is due on Sunday, March 24 at 11:59pm.
- Completed psets should be submitted to Gradescope.
- **Exercises** are for your own review only. They do not need to be submitted and will not be graded.
- **Complete all problems 1–3 and *one of either* 4 or 5.**

1 Exercises

Exercise 1

Proposition 1. *Let V be a Euclidean space and let*

$$\|v\| \coloneqq \sqrt{\langle v, v\rangle}.$$

Then, $\|v\|$ is a norm.

Proof. We need to prove positivity, homogeneity, and the triangle inequality.

Let $v \in V$. First, $\langle v, v\rangle \geq 0$ by definition of an inner product where $\langle v, v\rangle = 0$ iff $v = 0$. We deduce that $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$.

Next, let $v \in V$ and $\alpha \in \mathbb{R}$. Then,

$$\|\alpha v\| = \sqrt{\langle \alpha v, \alpha v\rangle} = \sqrt{\alpha^2 \langle v, v\rangle} = |\alpha| \sqrt{\langle v, v\rangle} = |\alpha| \|v\|.$$

The second equality follows from bilinearity of the inner product.

Finally, let $u, v \in V$. Our goal is to prove that $\|u + v\| \leq \|u\| + \|v\|$. Consider $f(\alpha, \beta) \coloneqq \|\alpha u + \beta v\|^2$. This is nonnegative for all $\alpha, \beta \in \mathbb{R}$. Additionally,

$$f(\alpha, \beta) = \alpha^2 \|u\|^2 + 2\alpha\beta \langle u, v\rangle + \beta^2 \|v\|^2 = \begin{pmatrix}\alpha \\ \beta\end{pmatrix}^\intercal \begin{pmatrix}\|u\|^2 & \langle u, v\rangle \\ \langle u, v\rangle & \|v\|^2\end{pmatrix} \begin{pmatrix}\alpha \\ \beta\end{pmatrix}.$$

As $f(\alpha, \beta)$ is nonnegative for all α, β, we deduce that the 2×2 matrix on the right is positive semidefinite and that its determinant must be nonnegative. Thus,

$$\|u\|^2 \|v\|^2 - \langle u, v\rangle^2 \geq 0.$$

Rearranging, we get the *Cauchy-Schwarz Inequality*:

$$|\langle u, v\rangle| \leq \|u\| \|v\| .$$

Now, we have that

$$\begin{aligned}(\|u\| + \|v\|)^2 &= \|u\|^2 + 2\|u\| \|v\| + \|v\|^2 \\ &\geq \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \\ &= \|u + v\|^2 .\end{aligned}$$

■

Exercise 2

Proposition 2. *Let $p > 0$. For $x \in \mathbb{R}^n$, define*

$$\|x\|_p \coloneqq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} .$$

If $p \in (0, 1)$ and $n \geq 2$, then $\|x\|_p$ is not *a norm.*

Proof. It is not hard to see that homogeneity and nonnegativity hold for any $p > 0$. We will prove that the triangle inequality does *not* hold for $p \in (0, 1)$ and $n \geq 2$. Let $p \in (0, 1)$. We compute:

$$\|e_1 + e_2\|_p = 2^{1/p} > 2 = \|e_1\|_p + \|e_2\|_p .$$

■

Exercise 3

Proposition 3. *The affine image of a convex set is a convex set.*

Proof. Let V, W be Euclidean spaces and let $\mathcal{L} : V \to W$ be an affine transformation. That is, if $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ satisfy $\sum_{i=1}^{k} \alpha_i = 1$ and $v_i \in V$, then

$$\mathcal{L}\left(\sum_{i=1}^{k} \alpha_i v_i\right) = \sum_{i=1}^{k} \alpha_i \mathcal{L}(v_i).$$

Let $C \subseteq V$ be a convex set. Our goal is to show that $\mathcal{L}(C)$ is convex. Let $w_1, w_2 \in \mathcal{L}(C)$. By assumption, there exists $v_1, v_2 \in C$ so that $\mathcal{L}(v_i) = w_i$. As C is convex, for every $\lambda \in [0, 1]$,

$$\lambda v_1 + (1 - \lambda) v_2 \in C.$$

As $\mathcal{L}$ is affine, we have that

$$\begin{aligned}\mathcal{L}(C) &\ni \mathcal{L}(\lambda v_1 + (1 - \lambda) v_2) \\ &= \lambda \mathcal{L}(v_1) + (1 - \lambda) \mathcal{L}(v_2) \\ &= \lambda w_1 + (1 - \lambda) w_2.\end{aligned}$$

■

Exercise 4

Proposition 4. *Let $C \subseteq \mathbb{R}^n$ be a convex set. Let $x \in \operatorname{rint}(C)$ and $y \in \operatorname{cl}(C)$. Prove that for all $\theta \in [0,1)$, that $(1-\theta)x + \theta y \in \operatorname{rint}(C)$.*

Proof. Let $\epsilon > 0$ so that

$$B(x,\epsilon) \cap \operatorname{aff}(C) \subseteq C.$$

This exists as $x \in \operatorname{rint}(C)$.

Let $\delta > 0$ and let $\bar{y} \in C$ so that $\|\bar{y} - y\| \leq \delta$. This exists as $y \in \operatorname{cl}(C)$.

Now, as C is convex, we have that for all $\theta \in [0,1)$,

$$\begin{aligned} C &\supseteq B\left((1-\theta)x + \theta\bar{y}, (1-\theta)\epsilon\right) \\ &\supseteq B\left((1-\theta)x + \theta y, (1-\theta)\epsilon - \theta\delta\right). \end{aligned}$$

The radius of this set is positive for all $\theta \in [0, \frac{\epsilon}{\epsilon+\delta})$. Thus, $(1-\theta)x + \theta y \in \operatorname{rint}(C)$ for all $\theta \in [0, \frac{\epsilon}{\epsilon+\delta})$. Letting $\delta \to 0$ completes the proof. ■

2 Problems

Problem 1 [25pts]

Given $A \in \mathbb{S}^n$ and $B \in \mathbb{S}^m$, the Kronecker product $A \otimes B$ is the $\mathbb{S}^{mn}$ matrix given in block form as

$$A \otimes B = \begin{pmatrix} A_{1,1}B & \dots & A_{1,n}B \\ \vdots & \ddots & \vdots \\ A_{n,1}B & \dots & A_{n,n}B \end{pmatrix}$$

Proposition 5. *Suppose $A \in \mathbb{S}^n_+$ and $B \in \mathbb{S}^m_+$. Then, $A \otimes B \succeq 0$.*

Proof. Let $\alpha_1, \dots, \alpha_n \in \mathbb{R}_+$ and $a_1, \dots, a_n \in \mathbb{R}^n$ denote the eigenvalues and an eigenbasis of A. Similarly define $\beta_1, \dots, \beta_m$ and $b_1, \dots, b_m$.

For $i \in [n]$ and $j \in [m]$, define

$$\gamma_{ij} = \alpha_i \beta_j \geq 0 \qquad\qquad c_{ij} = \begin{pmatrix} (a_i)_1 b_j \\ \vdots \\ (a_i)_n b_j \end{pmatrix} \in \mathbb{R}^{nm}.$$

Our goal is to show that c_{ij} are an eigenbasis with associated eigenvalues $\gamma_{ij} \geq 0$.

First

$$
\begin{aligned}
(A \otimes B)c_{ij} &= \begin{pmatrix} (A_{1,1}(a_i)_1 + \cdots + A_{1,n}(a_i)_n)(Bb_j) \\ \vdots \\ (A_{n,1}(a_i)_1 + \cdots + A_{n,n}(a_i)_n)(Bb_j) \end{pmatrix} \\
&= \alpha_i \beta_j \begin{pmatrix} (a_i)_1 b_j \\ \vdots \\ (a_i)_n b_j \end{pmatrix} = \gamma_{ij} c_{ij}.
\end{aligned}
$$

Next,

$$
\begin{aligned}
\langle c_{ij}, c_{i'j'} \rangle &= \left\langle \begin{pmatrix} (a_i)_1 b_j \\ \vdots \\ (a_i)_n b_j \end{pmatrix}, \begin{pmatrix} (a_{i'})_1 b_{j'} \\ \vdots \\ (a_{i'})_n b_{j'} \end{pmatrix} \right\rangle \\
&= \langle a_i, a_{i'} \rangle \langle b_j, b_{j'} \rangle .
\end{aligned}
$$

We see that $\langle c_{ij}, c_{i'j'} \rangle$ is positive if $(i, j) = (i', j')$ and zero otherwise. ■

Problem 2 [25pts]

Given a symmetric matrix $A \in \mathbb{S}^n$, let $\text{Inertia}(A) \coloneqq (n_-, n_0, n_+)$ denote the number of negative eigenvalues, number of zero eigenvalues, and number of positive eigenvalues of A.

Proposition 6. *Let $A \in \mathbb{S}^n$ and let $P \in \mathbb{R}^{n \times n}$ be invertible. Then,*

$$
\textit{Inertia}(A) = \textit{Inertia}(P^\intercal A P).
$$

Proof. Let $(n_-(A), n_0(A), n_+(A))$ denote the inertia of A.

First, let $V \coloneqq \ker(A)$. This is a subspace of $\mathbb{R}^n$ of dimension $n_0(A)$. Now, set $W \coloneqq P^{-1}V$. This again has dimension $n_0(A)$. Then, for any $w = P^{-1}v \in W$ we have

$$
(P^\intercal A P)w = P^\intercal A v = 0.
$$

Thus, $n_0(B) \geq n_0(A)$.

Next, let V denote the subspace spanned by the eigenvectors of A corresponding to positive eigenvalues. This is a subspace of $\mathbb{R}^n$ of dimension $n_+(A)$. Now, set $W \coloneqq P^{-1}V$. This again has dimension $n_+(A)$. Then, for any $w = P^{-1}v \in W \setminus \{0\}$ we have

$$
w(P^\intercal A P)w = v^\intercal A v > 0.
$$

By Courant-Fischer Theorem, $n_+(B) \geq n_+(A)$.

An analogous proof shows that $n_-(B) \geq n_-(A)$.

Thus, we have shown that

$$
n_-(A) \leq n_-(B) \qquad n_0(A) \leq n_0(B) \qquad n_+(A) \leq n_+(B).
$$

Reversing the roles of A and B completes the proof.[1] ■

[1] Or, simply observing that $n_- + n_0 + n_+ = n$.

Problem 3 [25pts]

Proposition 7. *The nonnegative orthant is self-dual.*

Proof. Let $\mathcal{K} := \mathbb{R}^n_+$ denote the nonnegative orthant.

We first show that $\mathcal{K}_* \subseteq \mathcal{K}$: Suppose $x \notin \mathcal{K}$. That is, there is some coordinate, say i, so that $x_i < 0$. Then, $e_i \in \mathcal{K}$ satisfies

$$\langle e_i, x\rangle < 0.$$

We deduce that $x \notin \mathcal{K}_*$.

Next, we show that $\mathcal{K} \subseteq \mathcal{K}_*$. That is, given $x, y \in \mathcal{K}$, we need to show that

$$\langle x, y\rangle \geq 0.$$

The left-hand side is just a sum of n nonnegative terms. ■

Proposition 8. *The second-order cone is self-dual.*

Proof. Let $\mathcal{K} := \mathcal{L}^{1+n}$ denote the second-order cone.

We first show that $\mathcal{K}_* \subseteq \mathcal{K}$: Suppose $(t, x) \notin \mathcal{K}$. That is, $t < \|x\|_2$. We will show that $(t, x) \notin \mathcal{K}_*$. If $t < 0$, then clearly $(t, x) \notin \mathcal{K}_*$ as $(1, 0) \in \mathcal{K}$ and $\langle (t, x), (1, 0)\rangle = t$. Now assume $t \geq 0$ and $t < \|x\|_2$. Thus, $x \neq 0$. Define $(s, y) \in \mathcal{K}$ by setting $s = \|x\|_2$ and $y = -x$. Then,

$$\left\langle \begin{pmatrix} s \\ y \end{pmatrix}, \begin{pmatrix} t \\ x \end{pmatrix} \right\rangle = st + \langle x, y\rangle < \|x\|_2^2 - \|x\|_2^2.$$

We deduce that $(t, x) \notin \mathcal{K}_*$.

Next, we show that $\mathcal{K} \subseteq \mathcal{K}_*$. That is, given $(t, x), (s, y) \in \mathcal{L}^{1+n}$, need to show that

$$\langle (t, x), (s, y)\rangle \geq 0.$$

We compute

$$\begin{aligned} \langle (t, x), (s, y)\rangle &= st + \langle x, y\rangle \\ &\geq st - \|x\|_2 \|y\|_2 \\ &\geq st - st. \end{aligned}$$

The first inequality follows from Cauchy-Schwarz and the second inequality follows from the definition of the second-order cone. ■

Proposition 9. *The semidefinite cone is self-dual.*

Proof. Let $\mathcal{K} := \mathbb{S}^n_+$ denote the positive semidefinite cone.

We first show that $\mathcal{K}_* \subseteq \mathcal{K}$: Suppose $X \notin \mathcal{K}$. That is, there is some $v \in \mathbb{R}^n$ so that $v^\intercal A v < 0$. We can rewrite this as

$$\langle A, vv^\intercal\rangle = v^\intercal A v < 0$$

Note that $vv^\intercal \in \mathcal{K}$. Thus, $A \notin \mathcal{K}_*$.

Next, we show that $\mathcal{K} \subseteq \mathcal{K}_*$. That is, given $X, Y \in \mathcal{K}$, we need to show that $\langle X, Y \rangle \geq 0$. By the spectral theorem, we may write $X = \sum_{i=1}^n \lambda_i v_i v_i^\intercal$, where $\lambda_i \geq 0$. Then,

$$\begin{aligned} \langle X, Y \rangle &= \left\langle \sum_{i=1}^n \lambda_i v_i v_i^\intercal, Y \right\rangle \\ &= \sum_{i=1}^n \lambda_i (v_i^\intercal Y v_i). \end{aligned}$$

As $Y \succeq 0$, the term in parentheses in the last line is nonnegative for all i. We conclude that $\mathcal{K} \subseteq \mathcal{K}_*$. ■

Problem 4 [25pts]

In sparse recovery, the goal is to recover a sparse vector $x^\star \in \mathbb{R}^n$ given linear measurements $(A, b) \in \mathbb{R}^{m \times n} \times \mathbb{R}^m$ where $b = Ax^\star$. A convex-optimization approach to this problem is to output the optimizer of

$$\min_{x \in \mathbb{R}^n} \left\{ \|x\|_1 : Ax = b \right\}.$$

Given a subset $S \subseteq [n]$ and a vector $x \in \mathbb{R}^n$, let x_S denote the restriction of x onto the set S. Let S^c denote the complement of S.

Proposition 10. *Let $S^\star$ denote the support of $x^\star$. Then, the descent cone at $x^\star$ is*

$$\left\{ \delta \in \mathbb{R}^n : \begin{array}{l} \delta \in \ker(A) \\ \langle \operatorname{sign}(x^\star), \delta_{S^\star} \rangle + \left\| \delta_{(S^\star)^C} \right\|_1 \leq 0 \end{array} \right\}$$

Proof. Let $\mathcal{K}$ denote the descent cone at $x^\star$. By definition,

$$\mathcal{K} := \left\{ \delta \in \mathbb{R}^n : \begin{array}{l} \forall \epsilon > 0 \text{ small enough} : \\ x^\star + \epsilon\delta \text{ is feasible} \\ \text{obj. value at } x^\star + \epsilon\delta \leq \text{obj. value at } x^\star \end{array} \right\}.$$

We specialize this to the current problem:

$$\mathcal{K} = \left\{ \delta \in \mathbb{R}^n : \begin{array}{l} \forall \epsilon > 0 \text{ small enough} : \\ Ax^\star + \epsilon A\delta = b \\ \|x^\star + \epsilon\delta\|_1 \leq \|x^\star\|_1 \end{array} \right\}.$$

Recall that $Ax^\star = b$. Let $S^\star$ denote the support of $x^\star$. We can decompose

$$\|x^\star + \epsilon\delta\|_1 = \|x^\star + \epsilon\delta_{S^\star}\|_1 + \epsilon \left\| \delta_{(S^\star)^c} \right\|_1.$$

When $\epsilon > 0$ is small enough, the sign of $(x^\star + \epsilon\delta_{S^*})$ is the same as the sign of $x^\star$, thus for all $\epsilon > 0$ small enough,

$$\|x^\star + \epsilon\delta\|_1 = \|x^\star\|_1 + \epsilon \langle \operatorname{sign}(x^\star), \delta_{S^\star} \rangle + \epsilon \left\| \delta_{(S^\star)^c} \right\|_1.$$

This gives us the form of $\mathcal{K}$ we set out to prove. ■

The matrix A is said to satisfy the *nullspace property at order k* if for all sets $S \subseteq [n]$ with $|S| \leq k$ and for all $\delta \in \ker(A) \setminus \{0\}$, we have

$$\|\delta_S\|_1 < \|\delta_{S^c}\|_1 .$$

Proposition 11. *Suppose $x^\star$ is $\leq k$-sparse and A satisfies the nullspace property at order k. Then, the descent cone at $x^\star$ is trivial.*

Proof. By the previous result, we have that any δ in the descent cone at $x^\star$ satisfies

$$\delta \in \ker(A) \qquad \text{and} \qquad \langle \operatorname{sign}(x^\star), \delta_{S^\star} \rangle + \left\|\delta_{(S^\star)^c}\right\|_1 \leq 0.$$

We can apply the nullspace property at order k to the set $S^\star$ to deduce that either $\delta = 0$ or

$$|\langle \operatorname{sign}(x^\star), \delta_{S^\star} \rangle| \leq \|\delta_{S^*}\| < \left\|\delta_{(S^*)^c}\right\| ,$$

a contradiction. We conclude that $\delta = 0$ and that the descent cone at $x^\star$ is trivial. ■

Proposition 12. *If A does not satisfy the nullspace property at order k, then there exists a k-sparse $x^\star$ for which the convex-optimization approach may fail to recover $x^\star$.*

Proof. By definition, there exists a set of coordinates S of size at most k and a

$$\delta \in \ker(A) \setminus \{0\} : \qquad \|\delta_S\|_1 \geq \|\delta_{S^c}\|_1 .$$

Set $x^\star = -\operatorname{sign}(\delta_S)$. This is a $\leq k$ sparse vector. We verify that δ is in the descent cone at $x^\star$:

$$\langle \operatorname{sign}(x^\star), \delta_S \rangle + \|\delta_{S^c}\|_1 = -\|\delta_S\|_1 + \|\delta_{S^c}\|_1 \leq 0.$$

We conclude that $x^\star$ is not the unique optimal solution to this problem (or even necessarily *an* optimal solution). ■

Problem 5 [25pts]

Given a permutation σ of $[n]$, we can associate σ with the $n \times n$ permutation matrix

$$(X^\sigma)_{i,j} = \begin{cases} 1 & \text{if } \sigma(i) = j \\ 0 & \text{else} \end{cases} .$$

Let $\mathcal{P}_n$ denote the set of all $n!$ permutation matrices.

Define the set of doubly stochastic matrices:

$$\mathrm{DS}_n \coloneqq \left\{ X \in \mathbb{R}^{n \times n} : \begin{array}{c} X \geq 0 \\ X^\intercal 1_n = 1_n \\ X 1_n = 1_n \end{array} \right\} .$$

Proposition 13. *Let* $n \geq 1$*, then* $\mathrm{conv}(\mathcal{P}_n) = \mathrm{DS}_n$*.*

Proof. First, we check that $\mathrm{conv}(\mathcal{P}_n) \subseteq \mathrm{DS}_n$: Let $X \in \mathcal{P}_n$. Then, X is a binary matrix where each row and column contains exactly one 1. We deduce that $X \in \mathrm{DS}_n$. As DS_n is convex and $\mathrm{conv}(\mathcal{P}_n)$ is the smallest convex set containing $\mathcal{P}_n$, we conclude that $\mathrm{conv}(\mathcal{P}_n) \subseteq \mathrm{DS}_n$.

Now, we check that $\mathrm{DS}_n \subseteq \mathrm{conv}(\mathcal{P}_n)$: Let $Y \in \mathrm{DS}_n$. We will prove that $Y \in \mathrm{conv}(\mathcal{P}_n)$ via induction on the number of nonzero entries in Y. Y must have some number of nonzero entries between n and n^2 (the lower bound comes from the fact that each row and column needs to have at least one nonzero entry).

If Y has exactly n nonzero entries, then we are done because Y will itself be a permutation matrix so $Y \in \mathcal{P}_n$.

Now, suppose Y has $> n$ nonzero entries.

Our first step is to deduce that the support of Y contains a permutation matrix. We will do this by applying Hall's marriage theorem to the following bipartite graph: Let the left vertex set be indexed by $i \in [n]$ and right vertex set be indexed by $j \in [n]$ and connect (i, j) with an edge if $Y_{i,j} > 0$. In order to apply Hall's marriage theorem, we must check that for any subset $\mathcal{L}$ of the left vertices that the number of neighbors of $\mathcal{L}$ is at least $|\mathcal{L}|$. Let $\mathcal{N}(\mathcal{L})$ denote the neighbors of $\mathcal{L}$. To see that this is true, let $\mathcal{L}$ be any subset of the left vertices. As $Y \in \mathrm{DS}_n$, it holds that

$$\begin{aligned}
|\mathcal{L}| &= \sum_{i\in\mathcal{L}} \sum_{j=1}^{n} Y_{i,j} \\
&= \sum_{i\in\mathcal{L}} \sum_{j\in\mathcal{N}(\mathcal{L})} Y_{i,j} \\
&= \sum_{j\in\mathcal{N}(\mathcal{L})} \sum_{i\in\mathcal{L}} Y_{i,j} \\
&\leq |\mathcal{N}(\mathcal{L})| .
\end{aligned}$$

Thus, we may apply Hall's marriage theorem to conclude that the support of Y contains a permutation matrix X.

Now, set $\theta > 0$ so that $Y - \theta X \geq 0$ has support strictly less than the support of Y. By induction, we have that

$$\frac{1}{1-\theta}(Y - \theta X) \in \mathrm{conv}(\mathcal{P}_n).$$

Then, as $X \in \mathcal{P}_n$, we conclude that

$$Y = \theta X + (1-\theta)\left(\frac{1}{1-\theta}(Y - \theta X)\right) \in \mathrm{conv}(\mathcal{P}_n).$$

■