DSA3102: Solutions to Problem Set -1
Assigned: 17/08/23
Never Due

1. Let $p \geq 1$. Show that the dual norm of $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}_+$ is $\|\cdot\|_q$ where $\frac{1}{p} + \frac{1}{q} = 1$. Hence, show that the dual of the dual norm is the original norm.

   Hint You may need Hölder's inequality. Find out what this is.

   Solution: The dual norm $\|\cdot\|_*$ is defined as

   $$\|z\|_* := \sup\{z^T x : x \in \mathbb{R}^n, \|x\| \leq 1\}$$

   for all $z \in \mathbb{R}^n$. Fix $1 < p, q < \infty$ where $\frac{1}{p} + \frac{1}{q} = 1$. Fix $z = (z_1, \ldots, z_n)$. We will show that

   $$\sup\left\{\sum_{i=1}^n z_i x_i : x = (x_1, \ldots, x_n) \in \mathbb{R}^n : \|x\|_q \leq 1\right\} = \|z\|_p.$$

   Assume without loss of generality that $z \neq 0$ otherwise both sides are zero. We have by Hölder's inequality that

   $$\sum_{i=1}^n z_i x_i \leq \sum_{i=1}^n |z_i x_i| \leq \|z\|_p \|x\|_q \leq \|z\|_p.$$

   Hence maximizing over all x yields the inequality $\leq$.

   Next we construct a vector y that achieves the bound with equality. We put

   $$x_i := \text{sign}(z_i)|z_i|^{p-1}, \qquad \forall\, i = 1, \ldots, n$$

   We then calculate

   $$\sum_{i=1}^n z_i x_i = \sum_{i=1}^n z_i \text{sign}(z_i)|z_i|^{p-1} = \sum_{i=1}^n |z_i|^p = \|z\|_p^p.$$

   Furthermore,

   $$\|x\|_q^q = \sum_{i=1}^n |x_i|^q = \sum_{i=1}^n |\text{sign}(z_i)|z_i|^{p-1}|^q = \sum_{i=1}^n |z_i|^{q(p-1)} = \sum_{i=1}^n |z_i|^p = \|z\|_p^p.$$

   where here we used that $\frac{1}{p} + \frac{1}{q} = 1$ so $q(p-1) = p$. Now choose

   $$y := \frac{x}{\|x\|_q}$$

   where here we used the fact that $z \neq 0$ so $\|x\|_q \neq 0$. By construction $\|y\|_q = 1$ and

   $$\sum_{i=1}^n z_i y_i = \frac{1}{\|x\|_q} \sum_{i=1}^n z_i x_i.$$

Furthermore,

$$\frac{1}{\|x\|_q}\sum_{i=1}^{n} z_i x_i = \frac{1}{\|z\|_p^{p/q}}\sum_{i=1}^{n} z_i x_i = \frac{1}{\|z\|_p^{p/q}}\|z\|_p^p = \|z\|_p^{p-p/q} = \|z\|_p$$

where we used the fact that $p - p/q = 1$. Thus, we have found a y with $\|y\|_q \leq 1$ and $\sum_{i=1}^n z_i y_i = \|z\|_p$ as desired.

The dual of $\|\cdot\|_p$ is $\|\cdot\|_{p'}$ where $\frac{1}{p} + \frac{1}{p'} = 1$. Since p and p' are symmetric, the dual of $\|\cdot\|_{p'}$ is $\|\cdot\|_p$ so the dual of the dual norm is the original norm.

2. Let $A \in \mathbb{R}^{m\times n}$ be a matrix. Write down the definitions of the range $\mathcal{R}(A)$ and the nullspace $\mathcal{N}(A)$ of A. For a subspace $\mathcal{V} \subset \mathbb{R}^n$, write down the definition of the orthogonal complement $\mathcal{V}^\perp$. Show that

$$\mathcal{N}(A) = \mathcal{R}(A^T)^\perp.$$

Solution: The range is

$$\mathcal{R}(A) = \{Ax : x \in \mathbb{R}^n\}.$$

The nullspace is

$$\mathcal{N}(A) = \{x : Ax = 0\}.$$

The orthogonal complement of a subspace $\mathcal{V}$ is the set

$$\mathcal{V}^\perp = \{x : z^T x = 0, \forall\, z \in \mathcal{V}\}.$$

Now, let $q \in \mathcal{N}(A)$. Then we have the following implications:

$$\begin{aligned} Aq &= 0 \\ \implies z^T Aq &= 0, \qquad \forall\, z \in \mathbb{R}^n \\ \implies (A^T z)^T q &= 0, \qquad \forall\, z \in \mathbb{R}^n \\ \implies y^T q &= 0, \qquad \forall\, y \in \mathcal{R}(A^T) \\ \implies q &\in \mathcal{R}(A^T)^\perp \end{aligned}$$

This implies that

$$\mathcal{N}(A) \subset \mathcal{R}(A^T)^\perp$$

In the other direction, take a vector $z \in \mathcal{R}(A^T)^\perp$. Then we have

$$\begin{aligned} y^T z &= 0, \qquad \forall\, y \in \mathcal{R}(A^T) \\ \implies (A^T x)^T z &= 0, \qquad \forall\, x \in \mathbb{R}^m \\ \implies x^T Az &= 0, \qquad \forall\, x \in \mathbb{R}^m \\ \implies Az &= 0, \qquad \forall\, z \in \mathcal{R}(A^T)^\perp \\ \implies z &\in \mathcal{N}(A) \end{aligned}$$

This implies that

$$\mathcal{R}(A^T)^\perp \subset \mathcal{N}(A)$$

which leads to

$$\mathcal{N}(A) = \mathcal{R}(A^T)^\perp$$

3. Show that $\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$.

Solution: Each column of AB is a linear combination of the columns of A, which implies that $\mathcal{R}(AB) \subset \mathcal{R}(A)$. Hence,

$$\dim(\mathcal{R}(AB)) \leq \dim(\mathcal{R}(A))$$

or equivalently

$$\text{rank}(AB) \leq \text{rank}(A)$$

Each row of AB is a combination of the rows of B so rowspace$(AB) \subset$ rowspace(B) but the dimension of the rowspace is the dimension of the column space which is equal to the rank so

$$\text{rank}(AB) \leq \text{rank}(B)$$

as desired.

4. Let $A \in \mathbf{S}^n$ (where recall that $\mathbf{S}^n$ is the set of all real symmetric $n \times n$ matrices) have eigen-decomposition $A = Q\Lambda Q^T$ where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Show that $\lambda_i(A), i \in \{1, \ldots, n\}$ are real. Show that eigenvectors of distinct eigenvalues are orthogonal.

   Solution: First we show that all eigenvalues must be real. For any eigenvector $u \neq 0$, we have

   $$Au = \lambda u$$

   where λ is the corresponding eigenvalue. Next, take the complex conjugate on both sides,

   $$A^* u^* = \lambda^* u^*$$

   But A is real so

   $$Au^* = \lambda^* u^*$$

   Next we premultiply the first equation by $(u^*)^T$, yielding

   $$(u^*)^T(Au) = (u^*)^T(\lambda u) = \lambda (u^*)^T u$$

   Furthermore, we have

   $$(u^*)^T(Au) = (A^T u^*)^T u = (Au^*)^T u = \lambda^* (u^*)^T u$$

   Combining the above equations yields

   $$\lambda^* (u^*)^T u = \lambda (u^*)^T u$$

   Since eigenvectors are non-zero, we have $\lambda^* = \lambda$ so λ is real as desired.

   Let λ and $\tilde{\lambda}$ be distinct eigenvalues, i.e., $\lambda \neq \tilde{\lambda}$. We have

   $$Au = \lambda u, \qquad A\tilde{u} = \tilde{\lambda}\tilde{u}.$$

   Premultiplying the first equation by $\tilde{u}^T$, we obtain

   $$\lambda \tilde{u}^T u = \tilde{u}^T A u = (A^T \tilde{u})^T u = (A\tilde{u})^T u = (\tilde{\lambda}\tilde{u})^T u = \tilde{\lambda}\tilde{u}^T u$$

   Thus, we have

   $$(\lambda - \tilde{\lambda})\tilde{u}^T u = 0$$

   Since $\lambda \neq \tilde{\lambda}$, we have $\tilde{u}^T u = 0$, i.e., $\tilde{u}$ and u are orthogonal as desired.

5. Let $A \in \mathbb{R}^{n \times n}$ be a matrix. Consider the linear system (fixed point equation)

   $$x^{(k+1)} = Ax^{(k)}.$$

   Let $x^{(0)} \in \mathbb{R}^n$ be the initial starting vector. Under what conditions on A does $x^{(k)}$ converge to a limit? What is the limit?

   Hint: $x^{(k)} = A^k x^{(0)}$. *Consider the eigen-decomposition of* A.

Solution: Let A have the eigen-decomposition

$$A = UDU^{-1}$$

Then, by using the hint, we obtain

$$x^{(k)} = UD^kU^{-1}x^{(0)}$$

because $A^k = UD^kU^{-1}$ through direct calculation. This is equivalent to

$$y^{(k)} = D^k y^{(0)}$$

if we define

$$y^{(j)} = U^{-1}x^{(j)}, \qquad \forall\, j \in \mathbb{N}$$

Note that D is a diagonal matrix and so

$$D^k = \operatorname{diag}(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k).$$

The elements of D converge to zero if and only if $|\lambda_i(A)| < 1$, i.e.,

$$\max_{1\le i\le n} |\lambda_i(A)| < 1, \qquad \Leftrightarrow \qquad \|D^k\|_F \to 0$$

Consequently, $\|y^{(k)}\|_2 \to 0$ if and only if $\max_{1\le i\le n} |\lambda_i(A)| < 1$. But $\|y^{(k)}\|_2 \to 0$ if and only if $\|x^{(k)}\|_2 \to 0$. Thus for the linear system to converge, it is necessary and sufficient that

$$\max_{1\le i\le n} |\lambda_i(A)| < 1$$

The limit is zero.

6. BV Problem 2.1

   Solution: This is readily shown by induction from the definition of convex set. We illustrate the idea for $k = 3$, leaving the general case to the reader. Suppose that $x, y, z \in C$ and $\theta_1 + \theta_2 + \theta_3 = 1$ with $\theta_j \ge 0$. We will show that $y = \sum_{j=1}^{3} \theta_j x_j \in C$. At least one of the θ_j is not equal to one; without loss of generality we can assume that $\theta_1 \ne 1$. Then we can write

   $$y = \theta_1 x_1 + (1-\theta_1)(\mu_2 x_2 + \mu_3 x_3)$$

   where

   $$\mu_2 = \frac{\theta_2}{1-\theta_1}, \quad \text{and} \quad \mu_2 = \frac{\theta_3}{1-\theta_1}$$

   Note that $\mu_2, \mu_3 \ge 0$ and $\mu_2 + \mu_3 = 1$ so by the convexity of C, we have that $\mu_2 x_2 + \mu_3 x_3 \in C$. Consequently, $y \in C$.

7. BV Problem 2.2

   Question: Show that a set is convex if and only if its intersection with any line is convex. Show that a set is affine if and only if its intersection with any line is affine.

   Solution: We prove the first part. The intersection of two convex sets is convex. Therefore if S is a convex set, the intersection of S with a line (which is convex) is also convex.

   Conversely, suppose the intersection of S with any line is convex. Take any two distinct points $x_1, x_2 \in S$. The intersection of S with the line through x_1 and x_2 is convex. Therefore convex combinations of x_1 and x_2 belong to the intersection, hence also to S.

   An argument roughly the same as the above also works for the affine case.

8. BV Problem 2.10

We will use the fact that a set is convex if and only if its intersection with an arbitrary line $L := \{\hat{x} + tv : t \in \mathbb{R}\}$ is convex. Let

$$C = \{x \in \mathbb{R}^n : x^T A x + b^T x + c \leq 0\}$$

where $A \in \mathbf{S}^n, b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

(a) We have

$$(\hat{x} + tv)^T A(\hat{x} + tv) + b^T(\hat{x} + tv) + c = \alpha t^2 + \beta t + \gamma$$

where

$$\alpha = v^T A c, \quad \beta = b^T v + 2\hat{x}^T A v, \quad \gamma = c + b^T \hat{x} + \hat{x} A^T \hat{x}$$

The intersection of C with the line defined by $\hat{x}$ and v is the set

$$\{\hat{x} + tv : \alpha t^2 + \beta t + \gamma \leq 0\}$$

which is convex if $\alpha \geq 0$. This is true for any v if $v^T A v \geq 0$, i.e., that $A \succeq 0$.

The converse does not hold. Take $A = -1, b = 0, c = -1$. Then A is not positive semidefinite but $C = \mathbb{R}$ is convex.

(b) Let $H = \{x : g^T x + h = 0\}$. We define α, β and γ as in the solution above. Additionally define

$$\delta = g^T v, \quad \epsilon = g^T \hat{x} + h$$

Without loss of generality we can assume that $\hat{x} \in H$, i.e., $\epsilon = 0$. The intersection of $C \cap H$ with the defined by $\hat{x}$ and v is

$$\{\hat{x} + tv : \alpha t^2 + \beta t + \gamma \leq 0, \delta t = 0\}$$

If $\delta = g^T v \neq 0$, the intersection is the singleton $\{\hat{x}\}$, if $\gamma \leq 0$, or it is empty. In either case, it is convex. If $\delta = 0$, the set reduces to

$$\{\hat{x} + tv : \alpha t^2 + \beta t + \gamma \leq 0\}$$

which is convex if $\alpha \geq 0$. Therefore $C \cap H$ is convex if

$$g^T v = 0 \quad \Rightarrow \quad v^T A v \geq 0$$

This is true if there exists λ such that $A + \lambda g g^T \geq 0$ because then

$$v^T A v = v^T (A + \lambda g g^T) v \geq 0$$

for all v satisfying $g^T v = 0$. Again, the converse is not true.

(c) Finally, we prove that a set S is convex $\Leftrightarrow$ its intersection with any line is convex. In the direction $\Rightarrow$, since S is convex and so is any line L, the intersection $S \cap L$ is convex. In the direction $\Leftarrow$, suppose S is a set such that $S \cap L$ is convex for all lines L. Take $x_1, x_2 \in S$. Consider the line L passing through x_1, x_2, i.e., $L = \{x : x = \theta x_1 + (1-\theta)x_2, \theta \in \mathbb{R}\}$. Since $S \cap L$ is convex, convex combinations $\theta x_1 + (1-\theta)x_2 \in S \cap L$ for $\theta \in [0,1]$. Clearly then $\theta x_1 + (1-\theta)x_2 \in S$ for all $\theta \in [0,1]$.

9. BV Problem 2.11

Solution: Assume that $\prod_i x_i \geq 1$ and $\prod_i y_i \geq 1$. Then consider the vector $z = \theta x + (1-\theta)y$. The product of its components is

$$\prod_i [\theta x_i + (1-\theta) y_i] \geq \prod_i x_i^\theta y_i^{1-\theta} = (\prod_i x_i)^\theta (\prod_i y_i)^{1-\theta} \geq 1$$

so the hyperbolic set is convex. We used the inequality

$$a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b$$

above.