HW1 Solutions

October 1,2019

1. We show that the four conditions are met:

   - $\|\cdot\|_1$ is nonnegative: $\|x\|_1 \geq 0$, for all $x \in \mathbb{R}^n$.
     $\|x\|_1 = \sum_{i=1}^n |x_i| \geq \sum_{i=1}^n 0 = 0, \ \forall x \in \mathbb{R}^n$.
   - f is definite: $\|x\|_1 = 0$ if and only if $x = 0$.
     $(\leftarrow)$if $\ x_i = 0, \ \forall i = 1, 2, ..., n \Rightarrow \|x\|_1 = \sum_{i=1}^n |x_i| = \sum_{i=1}^n 0 = 0$.
     $(\rightarrow)$if $x_i \neq 0 (i.e., x \neq 0) \Rightarrow \sum_{i=1}^n |x_i| > 0 \neq 0 \Rightarrow \sum_{i=1}^n |x_i| = 0$ only if $x = 0$.
   - $\|\cdot\|_1$ is homogeneous: $\|tx\|_1 = |t| \|x\|_1$, for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$.
     $\|tx\|_1 = \sum_{i=1}^n |tx_i| = \sum_{i=1}^n |t||x_i| = |t| \sum_{i=1}^n |x_i| = |t| \|x\|_1, \ \forall x \in \mathbb{R}^n, t \in \mathbb{R}$.
   - $\|\cdot\|_1$ satisfies the triangle inequality: $\|x+y\|_1 \leq \|x\|_1 + \|y\|_1$, for all $x, y \in \mathbb{R}^n$.
     $$\begin{aligned}\|x+y\|_1 &= \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n (|x_i| + |y_i|) = \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| \\ &= \|x\|_1 + \|y\|_1, \ \forall x, y \in \mathbb{R}^n.\end{aligned}$$

   Therefore, ℓ_1-norm is a norm.

2. We use the definition of operator norm on $\mathbb{R}^{m\times n}$
   $$\begin{aligned}\|X\|_{1,1} = \sup\{\|Xu\|_1 : \|u\|_1 \leq 1\} &= \sup\left\{\sum_{i=1}^m \left|\sum_{j=1}^n X_{i,j} u_j\right| : \sum_{j=1}^n |u_j| \leq 1\right\} \\ &\leq \sup\left\{\sum_{j=1}^n \sum_{i=1}^m |X_{i,j}| \, |u_j| : \sum_{j=1}^n |u_j| \leq 1\right\} \\ &\leq \sup\left\{\left[\max_j \left(\sum_{i=1}^m |X_{i,j}|\right)\right] \sum_{j=1}^n |u_j| : \sum_{j=1}^n |u_j| \leq 1\right\} \\ &\leq \sup\left\{\max_j \sum_{i=1}^m |X_{i,j}|\right\} \\ &= \max_j \sum_{i=1}^m |X_{i,j}| \,.\end{aligned}$$
   This can be achived by choosing u such that $u_j = 1$ for j that maximizes $\sum_{i=1}^m |X_{i,j}|$ and $u_j = 0$ otherwise.

3. $\|z\|_* = \sup\{z^T x : \|x\|_1 \leq 1\}$

$$\leq \sup\{\sum_{i=1}^{n} |z_i|\,|x_i| : \|x\|_1 \leq 1\}$$
$$= \sup\{\sum_{i=1}^{n} |z_i|\,|x_i| : \sum_{i=1}^{n} |x_i| \leq 1\}$$
$$\leq \sup\{\max_i |z_i| \sum_{i=1}^{n} |x_i| : \sum_{i=1}^{n} |x_i| \leq 1\}$$
$$\leq \sup\{\max_i |z_i|\}$$
$$= \max_i |z_i| = \|z\|_\infty .$$

This can be achived by choosing x such that $x_i = 1$ for i that maximizes $|z_i|$ and $x_i = 0$ otherwise.

4. (a) $\mathrm{tr}(AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}b_{ji} = \sum_{j=1}^n \sum_{i=1}^n b_{ji}a_{ij} = \mathrm{tr}(BA)$.

   (b) $\mathrm{tr}(tA + B) = \sum_{i=1}^n (ta_{ii} + b_{ii}) = t\sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = t \cdot \mathrm{tr}(A) + \mathrm{tr}(B)$.

5. $\|x - cy\|^2 = \langle x - cy, x - cy\rangle$

$$= \langle x, x\rangle + \langle -cy, x\rangle + \langle x, -cy\rangle + \langle -cy, -cy\rangle$$
$$= \langle x, x\rangle - c\langle y, x\rangle - c\langle x, y\rangle + c^2\langle y, y\rangle$$
$$= \langle x, x\rangle - 2c\langle x, y\rangle + c^2\langle y, y\rangle, \ \forall c \in \mathbb{R}.$$

Then we choose $c = \frac{\langle x,y\rangle}{\langle y,y\rangle}$.

$$\Rightarrow \|x - cy\|^2 = \langle x, x\rangle - 2\frac{\langle x, y\rangle}{\langle y, y\rangle}\langle x, y\rangle + \left(\frac{\langle x, y\rangle}{\langle y, y\rangle}\right)^2\langle y, y\rangle$$
$$= \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} \geq 0$$
$$\Rightarrow |\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle .$$

6. We show that the three conditions are met:

   - Conjugate symmetry: $\langle X, Y\rangle_F = \overline{\langle Y, X\rangle}_F$.
     $\overline{\langle Y, X\rangle}_F = \mathrm{tr}(\overline{Y^T X}) = \mathrm{tr}(Y^T X) = \mathrm{tr}(X^T Y) = \langle X, Y\rangle_F$.
   - Linearity in the first argument:
     $\langle cX, Y\rangle_F = c\langle X, Y\rangle_F$, and $\langle X + Z, Y\rangle_F = \langle X, Y\rangle_F + \langle Z, Y\rangle_F$.
     $\langle X + Z, Y\rangle_F = \mathrm{tr}((X + Z)^T Y)$
     $$= \mathrm{tr}((X^T + Z^T)Y)$$
     $$= \mathrm{tr}(X^T Y) + \mathrm{tr}(Z^T Y) \quad \text{(Problem by 4,(b))}$$
     $$= \langle X, Y\rangle_F + \langle Z, Y\rangle_F, \ \forall Z \in \mathbb{R}^{m\times n}.$$
     $\langle cX, Y\rangle_F = \mathrm{tr}(cX^T Y) = c \cdot \mathrm{tr}(X^T Y) = c\langle X, Y\rangle, \ \forall c \in \mathbb{R}.$ (Problem by 4,(b))

- Positive-definite: $\langle X, X\rangle_F > 0$, if $X \neq 0$.

  if $X \neq 0, \exists x_{i,j}$, then $\langle X, X\rangle_F = \text{tr}(X^T X) = \sum_{i=1}^m \sum_{j=1}^n x_{ij}^2 > 0$.

7. (a) $f(x) = \|Ax - b\|_2^2 = (Ax-b)^T(Ax-b) = (x^T A^T - b^T)(Ax-b)$

   $= x^T A^T Ax - b^T Ax - x^T A^T b + b^T b = x^T A^T Ax - 2b^T Ax + b^T b.$

   $\nabla f(x) = 2A^T Ax - 2A^T b = 2A^T(Ax - b).$

   (b) $\nabla^2 f(x) = \nabla_x(2A^T Ax - 2A^T b) = 2A^T A.$

8. (a) $\text{tr}(A) = \text{tr}(Q\Lambda Q^T) = \text{tr}(Q^T Q\Lambda) = \text{tr}(I\Lambda) = \text{tr}(\Lambda) = \sum_{i=1}^n \lambda_i.$

   (b) $(\rightarrow)$suppose that λ is an eigenvalue of A. Then there exists an nonzero eigenvector $x \in \mathbb{R}^n$ s.t., $Ax = \lambda x$. So $0 \leq x^T Ax = \lambda x^T x$. Since $x^T x$ is positive for all nonzero x,this implies that λ is non-negative.

   $(\leftarrow)$if $\lambda_i \geq 0 \quad \forall i = 1, 2, ..., n.$

   $\Rightarrow x^T Ax = x^T Q\Lambda Q^T x = z^T \Lambda z = \sum_{i=1}^n \lambda_i (z_i)^2 \geq 0 \ (let\ z = Q^T x).$

   (c) We set $z = Q^T x$.

   $\frac{x^T Ax}{x^T x} = \frac{x^T Q\Lambda Q^T x}{z^T QQ^T z} = \frac{z^T \Lambda z}{z^T z} = \frac{\lambda_{min}|z_1|^2 + ... + \lambda_{max}|z_n|^2}{|z_1|^2 + ... + |z_n|^2} \leq \frac{\lambda_{max}(|z_1|^2 + ... + |z_n|^2)}{|z_1|^2 + ... + |z_n|^2} = \lambda_{max}$

   $\Rightarrow x^T Ax \leq \lambda_{max} x^T x.$

   $\frac{x^T Ax}{x^T x} = \frac{z^T \Lambda z}{z^T z} = \frac{\lambda_{min}|z_1|^2 + ... + \lambda_{max}|z_n|^2}{|z_1|^2 + ... + |z_n|^2} \geq \frac{\lambda_{min}(|z_1|^2 + ... + |z_n|^2)}{|z_1|^2 + ... + |z_n|^2} \geq \lambda_{min} \Rightarrow x^T Ax \geq \lambda_{min} x^T x$

   $\Rightarrow \lambda_{min} x^T x \leq x^T Ax \leq \lambda_{max} x^T x.$