Optimization and Computational Linear Algebra for Data Science
Midterm solutions

October 22, 2019

Problem 1

(a) $E_1 = \{x \in \mathbb{R}^n \mid Ax = 3x\} = \mathrm{Ker}(A - 3\mathrm{Id}_n)$ is a subspace because a kernel of a matrix is a subspace.

(b) $E_2 = \{x \in \mathbb{R}^n \mid Ax = y\}$ is not a subspace because it does not contain the zero vector since $A0 = 0 \neq y$.

Problem 2 Statements **(a)**,**(c)**,**(d)**,**(f)** are equivalent to each other, and statements **(b)**,**(e)**,**(g)**,**(h)** are equivalent to each other.

Problem 3

(a) False. The following 4×3 has rank 3:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

(b) False. $\mathrm{rank}(\mathrm{Id}_n + \mathrm{Id}_n) = \mathrm{rank}(2\mathrm{Id}_n) = n \neq 2n = \mathrm{rank}(\mathrm{Id}_n) + \mathrm{rank}(\mathrm{Id}_n)$.

(c) True. If for all $u \in \mathbb{R}^n$, $\langle u, x \rangle = 0$ then in particular $\langle x, x \rangle = 0$ which gives $\|x\|^2 = 0$ and $x = 0$.

(d) False. Let $A = \mathrm{Diag}(1, 2, 3, \dots, n)$. $v_1 = e_1$ is an eigenvector associated to $\lambda_1 = 1$ and $v_2 = e_2$ is an eigenvector associated to $\lambda_2 = 2$. However $A(v_1 + v_2) = e_1 + 2e_2 \neq 3(e_1 + e_2)$: $(e_1 + e_2)$ is not an eigenvector of A associated with the eigenvalue $\lambda_1 + \lambda_2$.

Problem 4 Let $x \in \mathbb{R}^n$. We know that $P_S x$ is orthogonal to $x - P_S x$. Hence, by Pythagorean theorem:

$$\|x\|^2 = \|P_S x\|^2 + \|x - P_S x\|^2.$$

Problem 6 M is symmetric so by the spectral theorem there exists $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$, where $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, and P an orthogonal matrix such that

$$M = PDP^\mathsf{T}.$$

$M^{2019} = PDP^\mathsf{T}PDP^\mathsf{T} \cdots PDP^\mathsf{T} = PD^{2019}P^\mathsf{T}$, because $P^\mathsf{T}P = \mathrm{Id}_n$ (recall that P is orthogonal). We get that $PD^{2019}P^\mathsf{T} = 0$ which gives $0 = P^\mathsf{T}PD^{2019}P^\mathsf{T}P = D^{2019}$. Since $D^{2019} = \mathrm{Diag}(\lambda_1^{2019}, \dots, \lambda_n^{2019})$, we get that $\lambda_1 = \cdots = \lambda_n = 0$. We conclude that $M = P0P^\mathsf{T} = 0$.

Problem 6 We will prove that $\mathrm{rank}(A) = \mathrm{rank}(AA^\mathsf{T})$ which will imply the result, since the $n \times n$ matrix AA^T is invertible if and only if his rank equals n.

We have $\mathrm{Ker}(A^\mathsf{T}) = \mathrm{Ker}(AA^\mathsf{T})$. Indeed:

- if $x \in \mathrm{Ker}(A^\mathsf{T})$, then $AA^\mathsf{T}x = A0 = 0$. Hence $x \in \mathrm{Ker}(AA^\mathsf{T})$.
- if $x \in \mathrm{Ker}(AA^\mathsf{T})$, then $AA^\mathsf{T}x = 0$. Hence $\|A^\mathsf{T}x\|^2 = x^\mathsf{T}AA^\mathsf{T}x = 0$, which implies $A^\mathsf{T}x = 0$: $x \in \mathrm{Ker}(A^\mathsf{T})$.

Applying the rank-nullity theorem twice: $\mathrm{rank}(A^\mathsf{T}) = n - \dim(\mathrm{Ker}(A^\mathsf{T})) = n - \dim\mathrm{Ker}(AA^\mathsf{T}) = \mathrm{rank}(AA^\mathsf{T})$. We conclude that $\mathrm{rank}(A) = \mathrm{rank}(A^\mathsf{T}) = \mathrm{rank}(AA^\mathsf{T})$.

Problem 7

(a) $\text{rank}(A) \leq \min(n, m) = n$. Hence by the rank-nullity theorem we get that $\dim\text{Ker}(A) = m - \text{rank}(A) \geq m - n > 0$. Hence, we can find a non-zero vector $u \in \text{Ker}(A)$. Let $x = x^\star + u \neq x^\star$. We have

$$f(x) = \|Ax^\star + Au - y\| = \|Ax^\star - y\| = f(x^\star).$$

$x^\star$ is therefore not the unique minimizer of f.

(b) $P_{\text{Im}(A)}y$ belongs to $\text{Im}(A)$, hence there exists $x \in \mathbb{R}^m$ such that $P_{\text{Im}(A)}y = Ax$. By contradiction, assume that $Ax^\star \neq P_{\text{Im}(A)}y$. Then, by definition of $P_{\text{Im}(A)}y = Ax$,

$$\|Ax - y\| < \|Ax^\star - y\|.$$

This contradicts the fact that $x^\star$ minimizes f. Therefore $Ax^\star = P_{\text{Im}(A)}y$.

(c) Since $y - P_{\text{Im}(A)}y$ is orthogonal to $\text{Im}(A)$, we get that for all $x \in \mathbb{R}^m$,

$$\langle Ax, y - Ax^\star \rangle = 0,$$

because $Ax \in \text{Im}(A)$ and $P_{\text{Im}(A)}y = Ax^\star$.

(d) We deduce that for all $x \in \mathbb{R}^m$:

$$x^\top (A^\top y - A^\top A x^\star) = 0.$$

The vector $A^\top y - A^\top A x^\star$ is therefore orthogonal to every vector of $\mathbb{R}^m$: it is equal to 0.

(e) $\text{rank}(A^\top A) \leq \text{rank}(A) \leq n < m$, hence the $m \times m$ matrix $A^\top A$ is not invertible.