CSE 521 HW5

Rohan Mukherjee

November 16, 2023

1. (a) We first claim that $PMP^T \in \mathbb{R}^{m\times m}$ is symmetric. Indeed,

$$(PMP^T)^T = (P^T)^T M^T P^T = PMP^T$$

We recall that symmetric $M \in \mathbb{R}^{n\times n}$ is PSD iff $x^T M x \geq 0$ for all $x \in \mathbb{R}^n$. Thus, given an arbitrary $x \in \mathbb{R}^m$,

$$x^T PMP^T x = (P^T x)^T M (P^T x) \geq 0$$

Since $P^T x \in \mathbb{R}^n$ is just another vector.

(b) I claim that if M has k positive eigenvalues, then $M + vv^T$ has at most $k+1$ positive eigenvalues for any v. Indeed, enumerate the eigenvalues of M as $\lambda_1, \dots, \lambda_n$, and the eigenvalues of $M + vv^T$ as $\beta_1, \dots, \beta_n$. We have the following inequality:

$$\lambda_n \leq \beta_n \leq \cdots \leq \lambda_1 \leq \beta_1$$

In particular, since $\lambda_{k+1} \leq 0$, we have $\beta_{k+2} \leq \lambda_{k+1} \leq 0$, and in general $\beta_i \leq 0$ for $i > k+2$. Thus there can be at most $k+1$ positive eigenvalues. My second claim is that if M has k positive eigenvalues, then $M - vv^T$ has at most k positive eigenvalues. This follows similarly as last time, since $\lambda_{k+1} \leq \beta_{k+1} \leq 0$. With these two simple facts we can finish the problem. Write $M = \sum_{i=1}^n \lambda_i u_i u_i^T$. Clearly

$$PMP^T = \sum_{i=1}^n \lambda_i (Pu_i)(Pu_i)^T = \sum_{i=1}^n \operatorname{sgn}(\lambda_i)(\sqrt{|\lambda_i|}Pu_i)(\sqrt{|\lambda_i|}Pu_i)^T.$$

Starting from the 0 matrix, which has 0 positive eigenvalues, we now add each of the matrices in the above sum one at a time. For each $\lambda_i > 0$, we can pick up at most 1 positive eigenvalue, and for each $\lambda_i < 0$, can pick up at most 0 eigenvalues. Since there are exactly k positive eigenvalues, we conclude that PMP^T has at most k positive eigenvalues. Forsooth, the claim is upon us.

2. (a) If σ_i are the singular values of A, we know that σ_i^2 are the eigenvalues of A^TA. What can we say about $(A^TA)^T(A^TA) = A^TAA^TA$? Clearly, if v_i is the eigenvector associated with σ_i^2, we have that

$$A^TAA^TAv_i = A^TA\sigma_i^2 v_i = \sigma_i^4 v_i$$

So the eigenvalues of A^TAA^TA are precisely the σ_i^4. Now,

$$\|A^TA\|_F^2 = \sum_{i=1}^n \lambda_i((A^TA)^T(A^TA)) = \sum_{i=1}^n \sigma_i^4 \leq \sigma_1^2 \sum_{i=1}^n \sigma_i^2 = \|A\| \cdot \|A\|_F$$

We remark that this also gives equality conditions. Equality holds iff every singular value is equal.

(b) First, $\|A\sigma\|^2 = \sigma^TA^TA\sigma$. Second, notice that

$$\mathbb{E}\left[\|A\sigma\|^2\right] = \mathbb{E}\left[\sum_{i=1}^n \langle a_i, \sigma\rangle^2\right] = \mathbb{E}\left[\sum_{i=1}^n \left(\sum_{j=1}^n a_j^i \sigma_j\right)^2\right]$$
$$= \sum_{i=1}^n \sum_{1\leq j,k\leq n} \mathbb{E}\left[a_j^i a_k^i \sigma_j \sigma_k\right]$$

Next notice that $\mathbb{E}[\sigma_i] = 0$ and that $\mathbb{E}\left[\sigma_i^2\right] = \frac{1}{2}\cdot 1^2 + \frac{1}{2}\cdot(-1)^2 = 1$. Thus, this equals,

$$\sum_{i=1}^n \sum_{j=1}^n (a_j^i)^2 = \|A\|_F^2$$

Now,

$$\begin{aligned}
\Pr\left[\left|\sigma^TA^TA\sigma - \mathbb{E}\left[\sigma^TA^TA\sigma\right]\right| > \varepsilon\mathbb{E}\left[\sigma^TA^TA\sigma\right]\right] &= \Pr\left[\left|\|A\sigma\|^2 - \|A\|_F^2\right| > \varepsilon\|A\|_F^2\right] \\
&\leq 2\exp\left(-c\frac{\varepsilon^2\|A\|_F^4}{\|A^TA\|_F^2}\right) \\
&\leq 2\exp\left(-c\frac{\varepsilon^2\|A\|_F^4}{\|A\|^2\|A\|_F^2}\right) \\
&= 2\exp\left(-c\frac{\varepsilon^2\|A\|_F^2}{\|A\|^2}\right)
\end{aligned}$$

Very nice.

(c) We see clearly that $d(\sigma, E) = \|\sigma - \Pi_E\sigma\|$, as was shown in class and can be easily seen geometrically. Thus, $d(\sigma, E) = \|(I - \Pi_E)\sigma\| = \|\Pi_E^{\perp}\sigma\|$. Finding $w_1, \ldots, w_d$ an orthonormal basis for E, we have that $\Pi_E = \sum_{i=1}^d w_i w_i^T$. Extending this to a basis of

$\mathbb{R}^n$ as $w_1, \ldots, w_d, v_1, \ldots, v_{n-d}$, we have that $\Pi_E^{\perp} = \sum_{i=1}^{n-d} v_i v_i^T$. Notice now that

$$\Big\| \sum_{i=1}^{n-d} v_i v_i^T \Big\|_F^2 = \mathrm{Tr}\left(\sum_{1 \le i, j \le n-d} v_i v_i^T v_j v_J^T \right) = \mathrm{Tr}\left(\sum_{i=1}^{n-d} v_i v_i^T \right) = \sum_{i=1}^{n} \mathrm{Tr}\Big(v_i v_i^T \Big)$$
$$= \sum_{i=1}^{n-d} \mathrm{Tr}\Big(v_i^T v_i \Big) = \sum_{i=1}^{n-d} \|v_i\|^2 = n - d$$

We also claim that $\|\Pi_E^{\perp}\| \le 1$. This follows since $\|\Pi_E^{\perp}\|$ equals the max eigenvalue of $\Pi_E^{\perp}(\Pi_E^{\perp})^T = \Pi_E^{\perp}$, and the max eigenvalue of a projection matrix is at most 1. We conclude that

$$\Pr\left[\left| \frac{d(\sigma, E)}{n-d} - 1 \right| > \varepsilon \right] = \Pr\left[\left| \frac{\|\Pi_E^{\perp}\sigma\|}{\|\Pi_E^{\perp}\|_F} - 1 \right| > \varepsilon \right] \le 2\exp\left(-c \frac{\varepsilon^2 \|\Pi_E^{\perp}\|_F^2}{\|\Pi_E^{\perp}\|} \right)$$
$$= 2\exp\left(-c \frac{\varepsilon^2 (n-d)}{\|\Pi_E^{\perp}\|} \right)$$
$$\le 2\exp\Big(-c\varepsilon^2 (n-d) \Big)$$