Yale University
Department of Statistics and Data Science

Statistics 365/565

Issued: 04/12/2021 **Due:** 04/27/2021

Notation: $[k] = \{1, 2, \dots, k\}$. For a matrix $A \in \mathbb{R}^{m \times n}$ we will let $A_{(i,\cdot)}$ denote the i^{th} row and $A_{(\cdot,j)}$ denote the j^{th} column. **Both will be treated as column vectors.**

1 PCA

Suppose that we have data with covariance matrix $\Sigma = vv^T + \sigma^2 I$ for $v \in \mathbb{R}^p$.

Problem 1: Prove that the largest singular value of Σ is $\|v\|_2^2 + \sigma^2$. Some suggested ways to approach this, but no need to do the problem like this.

Part 1: The SVD is written as $\Sigma = VDV^T = \sum_{i=1}^p d_i v_i v_i^T$, since Σ is symmetric and PSD so $d_i \geq 0$. Let's just guess that one of the singular vectors is $\tilde{v}_1 = \frac{v}{\|v\|_2}$. Then from that we know we can build a set of $p-1$ orthogonal vectors $\tilde{v}_i$ for $i \in [2, p]$. So, all of the $\tilde{v}_i$ are orthogonal unit vectors. Let's put them into a matrix $V \in \mathbb{R}^{p \times p}$ so that the i^{th} column is $\tilde{v}_i$. Show that $\sigma^2 I = \sigma^2 \sum_{i=1}^p \tilde{v}_i \tilde{v}_i^T$

Part 2: Using the above, now show that $vv^T + \sigma^2 I = \|v\|_2^2 \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=1}^p \tilde{v}_i \tilde{v}_i^T$.

Part 3: Now show that

$$\Sigma = (\|v\|_2^2 + \sigma^2)\tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=2}^p \tilde{v}_i \tilde{v}_i^T$$

To conclude that $\Sigma = VDV^T$ where $D_{11} = \|v\|_2^2 + \sigma^2$ and all other $D_{ii} = \sigma^2$.

Problem 2: Prove that the top left (and right since Σ is symmetric) singular vector is $v/\|v\|_2$.

Problem 3: Suppose that data $x_i = vy_i + w_i$ where $y_i \sim N(0, 1)$ and $w_i \sim N(0, \sigma^2 I)$ where we assume that y_i and w_i are uncorrelated. Show that the covariance of x_i is Σ above. In a few sentences explain intuitively why it makes sense for $\frac{v}{\|v\|_2}$ to correspond to the largest singular vector of Σ (that is the singular vector associated with the largest singular value). You can do so by connecting this interpretation to any of the interpretations presented in lecture.

2 SVD

Finding the singular values and vectors of a matrix is important as we have discussed. One famous method of doing so is via the power iteration method. To find the top left and right singular vectors as well as the singular value we can perform the following algorithm. Consider our matrix $X \in \mathbb{R}^{p \times n}$. For all problems below we take $X = USV^T = \sum_{j=1}^{p \wedge n} u_j v_j^T \sigma_j$ where u_j and v_j are the left (resp. right) singular vectors while σ_j are the singular values sorted in decreasing order.

Initialize $v \sim N(0, I_n)$ and $u \sim N(0, I_p)$.

- Let $u = Xv/\|Xv\|_2$
- Let $v = X^T u/\|X^T u\|_2$
- Repeat until convergence

Problem 4: Suppose that $X = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$ with $\sigma_1 > \sigma_2 > 0$ is the SVD of X. So the u_i and v_i are unit vectors and $u_1^T u_2 = 0$ and $v_1^T v_2 = 0$. Suppose that we have a unit norm v such that $\alpha_1 = v^T v_1 > 0$ and $\alpha_2 = v^T v_2 > 0$. Let $u = Xv/\|Xv\|_2$. Let $v' = X^T u/\|X^T u\|_2$. Let $u' = Xv'/\|Xv'\|_2$. Show that $[u']^T u_1 > u^T u_1$. That is, after one step of the power method, u' is now more aligned with u_1 than the initial estimate u was.

We show this in parts

Part 1: Show that the following way to produce u' is equivalent to the one written in the previous paragraph.

- $\tilde{u} = Xv$
- $\tilde{v}' = X^T \tilde{u}$
- $u' = X\tilde{v}'/\|X\tilde{v}'\|_2$

That is, we only have to do the renormalization at the last step.

Part 2: Show the following equalities.

$$u = \frac{u_1 + \frac{\alpha_2 \sigma_2}{\alpha_1 \sigma_1} u_2}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^2}{\alpha_1^2 \sigma_1^2}}}$$

$$u' = \frac{u_1 + \frac{\alpha_2 \sigma_2^3}{\alpha_1 \sigma_1^3} u_2}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^6}{\alpha_1^2 \sigma_1^6}}}$$

Part 3: Finally, show the result.

Extra credit: Show that for the SVD $X = USV^T$ we have

$$\sigma_1 = \max_{\|u\|_2 \le 1, \|v\|_2 \le 1} u^T X v$$