Assignment 0(Part II)

BY XIAONAN PENG

Qishi Intermediate Machine Learning

Problem 1. Suppose $A \in \mathbb{R}^{m \times n}$, last time we have proved that $r(A) = r(A^T A) = r(A A^T)$. Now we can go further to details of SVD.

a) Prove the #nonzero eigenvalues of $A^T A = r(A)$. Then #zero eigenvalues $= n - r(A)$.

b) Prove $A^T A$ and $A A^T$ share the same nonzero eigenvalues.(Can they have negative eigenvalues?)

c) Suppose $A^T A v_i = \sigma_i^2 v_i, i = 1, \ldots, n$ and $\sigma_i^2 \geqslant \sigma_{i+1}^2$. Prove $u_i = \frac{A v_i}{\sigma_i}$ is eigenvector of $A A^T$ corresponding to eigenvalue σ_i^2. Then we have $A v_i = \sigma_i u_i$.

d) Write the pesudo code of SVD, suppose you can use the function eigen(A), which returns all eigenvalues and eigenvector matrix.

Problem 2. Compute the following derivatives, A, B are constant matrices, X is variable matrix.

a) $\frac{\partial}{\partial \beta}\left(\frac{1}{2}\|X\beta - y\|_2^2 + \frac{\lambda}{2}\|\beta\|_2^2\right)$, ridge regression.

b) $\frac{\partial \operatorname{tr}(A X)}{\partial X}$

c) $\frac{\partial \operatorname{tr}(A X B)}{\partial X}$

d)$\frac{\partial \operatorname{tr}(A X^{-1} B)}{\partial X}$

Problem 3. Suppose $\mathcal{D} = \{X_1, \ldots, X_n\} \sim^{\text{iid}} N(\mu, \Sigma)$ which is a $n - d$ Gaussian, we have the log likelihood

$$l(\mu, \Sigma | \mathcal{D}) = \ln\left(\prod_{i=1}^{n} p(X_i)\right)$$

compute the following derivatives

a) $\frac{\partial}{\partial \mu} l(\mu, \Sigma | \mathcal{D})$

b) $\frac{\partial}{\partial \Sigma} l(\mu, \Sigma | \mathcal{D})$

c) Give the MLE estimator.

Problem 4. Suppose $x \sim N(\mu, \Sigma)$, a $n - d$ Gaussian, compute $\mathbb{E}(\|x\|_2^2)$

Problem 5. Read the following

a) https://ccjou.wordpress.com/2012/05/30/lagrange-乘數法/

b) https://ccjou.wordpress.com/2017/02/07/karush-kuhn-tucker-kkt-條件/

c) Read ISL chapter 1 to 2