EE 510: Mathematical Foundations of Machine Learning **Winter 2020**

Exercises 6

Pages: 4.1-4.29 *Instructor Name: John Lipor*

Exercise 1

Explain why items 1 and 2 are true on pg. 4.3.

Exercise 2

Assume you obtain N noisy measurements $y_i = w^T x_i + n_i$, where $w \in \mathbb{R}^D$, $x_i \in \mathbb{R}^D$ for $i = 1, \ldots, N$, and n_i is a scalar that represents measurement noise. Define the matrix X and vector y such that finding w can be written as a least-squares problem.

Exercise 3

Derive the gradient equation given on pg. 4.9.

Exercise 4

Let $A \in \mathbb{R}^{m \times n}$ with $m > n$. Name a condition on A such that the matrix $A^T A$ is invertible. In this case, the least-squares solution is $\hat{x} = (A^T A)^{-1} A^T y$.

Exercise 5

Is the SVD applied on pg. 4.10 a full or thin SVD?

Exercise 6

Verify that $\Sigma_r^{-1} U_r^T y$ gives the $\hat{z}$ defined in Eq. (4.6).

Exercise 7

Why can the estimate $\hat{x}$ be recovered as $\hat{x} = Vz$ on pg. 4.12?

Exercise 8

Why do we need $M \geq N$ for $A \in \mathbb{R}^{M \times N}$ to have linearly independent columns?

Exercise 9

What size is V_0 when computing the SVD of a tall, full-rank matrix A? A tall matrix is one with more rows than columns.

Exercise 10

Use the properties of the pseudoinverse to show that $A^+ = V_r \Sigma_r^{-1} U_r^T$.

Exercise 11

Complete the exercise given on pg. 4.19.

Exercise 12

Verify that $Ax = 0$ for $x \in \mathcal{N}(A)$.

Exercise 13

Write in words what is meant by the optimization problem given in Eq. (4.16) on pg. 4.29.

Exercise 14

Why is $A^+ y \in \mathcal{R}(V_r)$ on pg. 4.29?

Exercise 15

BV 12.1