1 Back to Basics: Linear Algebra

Let $X \in \mathbb{R}^{m \times n}$. We do not assume that X has full rank.

(a) Give the definition of the rowspace, columnspace, and nullspace of X.

(b) Check (write an informal proof for) the following facts:

  (a) The rowspace of X is the columnspace of $X^\top$, and vice versa.
  (b) The nullspace of X and the rowspace of X are orthogonal complements.
  (c) The nullspace of $X^\top X$ is the same as the nullspace of X. *Hint: if v is in the nullspace of $X^\top X$, then $v^\top X^\top X v = 0$.*
  (d) The columnspace and rowspace of $X^\top X$ are the same, and are equal to the rowspace of X. *Hint: Use the relationship between nullspace and rowspace.*

2 Probability Review

There are n archers all shooting at the same target (bulls-eye) of radius 1. Let the score for a particular archer be defined to be the distance away from the center (the lower the score, the better, and 0 is the optimal score). Each archer's score is independent of the others, and is distributed uniformly between 0 and 1. What is the expected value of the worst (highest) score?

(a) Define a random variable Z that corresponds with the worst (highest) score.

(b) Derive the Cumulative Distribution Function (CDF) of Z.

(c) Derive the Probability Density Function (PDF) of Z.

(d) Calculate the expected value of Z.

3 Vector Calculus

[1]

[1] Good resources for matrix calculus are:

- The Matrix Cookbook: `https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf`
- Wikipedia: `https://en.wikipedia.org/wiki/Matrix_calculus`
- Khan Academy:

Below, $\mathbf{x} \in \mathbb{R}^d$ means that $\mathbf{x}$ is a $d \times 1$ column vector with real-valued entries. Likewise, $\mathbf{A} \in \mathbb{R}^{d \times d}$ means that $\mathbf{A}$ is a $d \times d$ matrix with real-valued entries. In this course, we will by convention consider vectors to be column vectors.

Consider $\mathbf{x}, \mathbf{w} \in \mathbb{R}^d$ and $\mathbf{A} \in \mathbb{R}^{d \times d}$. In the following questions, $\frac{\partial}{\partial \mathbf{x}}$ denotes the derivative with respect to $\mathbf{x}$, while $\nabla_{\mathbf{x}}$ denotes the gradient with respect to $\mathbf{x}$. Recall that $\nabla_{\mathbf{x}} f = \left(\frac{\partial f}{\partial \mathbf{x}}\right)^\top$.

Derive the following derivatives.

(a) $\frac{\partial \mathbf{w}^\top \mathbf{x}}{\partial \mathbf{x}}$ and $\nabla_{\mathbf{x}}(\mathbf{w}^\top \mathbf{x})$

(b) $\frac{\partial (\mathbf{w}^\top \mathbf{A}\mathbf{x})}{\partial \mathbf{x}}$ and $\nabla_{\mathbf{x}}(\mathbf{w}^\top \mathbf{A}\mathbf{x})$

(c) $\frac{\partial (\mathbf{w}^\top \mathbf{A}\mathbf{x})}{\partial \mathbf{w}}$ and $\nabla_{\mathbf{w}}(\mathbf{w}^\top \mathbf{A}\mathbf{x})$

(d) $\frac{\partial (\mathbf{w}^\top \mathbf{A}\mathbf{x})}{\partial \mathbf{A}}$ and $\nabla_{\mathbf{A}}(\mathbf{w}^\top \mathbf{A}\mathbf{x})$

(e) $\frac{\partial (\mathbf{x}^\top \mathbf{A}\mathbf{x})}{\partial \mathbf{x}}$ and $\nabla_{\mathbf{x}}(\mathbf{x}^\top \mathbf{A}\mathbf{x})$

(f) $\nabla^2_{\mathbf{x}}(\mathbf{x}^\top \mathbf{A}\mathbf{x})$

---

https://www.khanacademy.org/math/multivariable-calculus/multivariable-derivatives

- YouTube: https://www.youtube.com/playlist?list=PLSQl0a2vh4HC5feHa6Rc5c0wbRTx56nF7.