Introduction To Machine Learning **Spring 2024**

Exercise 0

Sharon Rotgaizer srotgaizer@gmail.com

Not for submission

1 Linear Algebra

A matrix A over $\mathbb{R}$ is called positive semidefinite (PSD) is $\forall v : v^\top A v \geq 0$.

1. Let $X \in \mathbb{R}^d$ be a random vector. Show that the covariance matrix Σ (such that $\Sigma_{ij} = \text{Cov}(X_i, X_j)$) is PSD.
2. Show that if A, B are PSD, then also $\alpha A + \beta B$, where $\alpha, \beta \geq 0$.
3. Show that $X^\top X$ is PSD and prove its eigenvalues are real (over $\mathbb{R}$).
   Note: this matrix is also symmeytic.
4. Prove that if two features are linearly dependent, the sample covariance matrix will have determinant of 0 (singular matrix).

2 Calculus

We say about function that it is convex, if $\forall t \in [0, 1]$ and $\forall x, y \in \mathbb{R}$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

See example here

1. Prove that $f(x) = \max\{0, x\}$ is a convex function.
2. Let $\mathbb{E} = \sum_{i=1}^{n}(x_i - w)^2$ be the "squared error".
   Explain why that function has global minimum (w.r.t w) and find it.
3. Let $\mathbb{E} = \sum_{i=1}^{n}|x_i - w|$ be the "absolute error".
   Find w that minimizes that error. No need to prove convexity.
4. Let A be an $n \times n$ matrix. Prove that:

$$\frac{\partial x^\top A x}{\partial x} = (A + A^\top)x$$

3 Porbability

1. Let $x_1, \ldots, x_n$ be independent samples taken from the continuous uniform distribution $\text{Uni}(a, b)$. Show that $\hat{a}_{MLE} = \min\{x_1, \ldots, x_n\}$ and $\hat{b}_{MLE} = \max\{x_1, \ldots, x_n\}$.
2. Let $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^{n\times 1}$ and $X \sim N(\mu, \Sigma_X)$. Let $Y = AX + b$. Show that $\Sigma_Y = A\Sigma_X A^\top$.
3. Let $X \in \mathbb{R}^d$ be a random vector. Denote its mean as $\mu \in \mathbb{R}^{d\times 1}$ and covariance matrix with $\Sigma \in \mathbb{R}^{d\times d}$. Recall that $\Sigma = \mathbb{E}[(X-\mu)(X-\mu)^\top]$. Show that:
$$\Sigma = \mathbb{E}[XX^\top] - \mu\mu^\top$$
4. A particle starts a random walk from the origin, in $\mathbb{Z}^2$, for $n \geq 200$ steps. At each step, it moves right / left / up / down with equal probability for each direction. Let $P = (X, Y) \in \mathbb{Z}^2$ the point that the particle has reached. Find the distribution of P.

4 Python

In this course we are going to work with Google Colab. However, feel free to work with any IDE that you want for this assignment.

Save HW0.ipynb to your Google Drive or download and upload it to your drive. There you can work online on your code, share with your partner, without the need for any external libraries download or version syncing.