Homework 1

Tuesday, January 30, 2024 7:58 PM

Question 1

We must show $\mathbb{R}^D$ is closed under addition and scalar multiplication and that 0 is in the set.

Consider $\vec{x} = [x_1, \dots, x_D]^T$ and $\vec{y} = [y_1, \dots, y_D]^T$ both elements of $\mathbb{R}^D$

For $\alpha \in \mathbb{R}$

$$\alpha \vec{x} + \vec{y} = [\alpha x_1 + y_1, \dots, \alpha x_D + y_D]^T \in \mathbb{R}^D$$

Furthermore $\underbrace{[0, \dots, 0]^T}_{D \text{ elements}} \in \mathbb{R}^D$ hence

$\mathbb{R}^D$ is infact a subspace

Question 2.

a.) Consider $[-1] \in \mathbb{R}^1$. Then $[\sqrt{-1}] = [i] \notin \mathbb{R}^1$

b.) $\mathbb{C}^D$ is closed under elementwise square roots and is a vector space.

Question 3.

Suppose $\vec{u}_1, \dots, \vec{u}_R \in \mathbb{R}^D$. We will show that $\text{span}[\vec{u}_1, \vec{u}_2, \dots, \vec{u}_R]$ is a subspace.

Recall: $\text{span}[\vec{u}_1, \dots, \vec{u}_R] = \left\{ \sum_{i=1}^{R} a_i \vec{u}_i : a_i \in \mathbb{R} \right\}$

Let $\vec{x} = a_1 \vec{u}_1 + a_2 \vec{u}_2 + \dots + a_R \vec{u}_R$ and
$\vec{y} = b_1 \vec{u}_1 + b_2 \vec{u}_2 + \dots + b_R \vec{u}_R$

Then for some $\alpha \in \mathbb{R}$,

$$\alpha \vec{x} + \vec{y} = (\alpha a_1 + b_1)\vec{u}_1 + \dots + (\alpha a_R + b_R)\vec{u}_R$$

which is in the span as $\alpha a_i + b_i \in \mathbb{R}$ for all i.

Furthermore $\vec{0} = 0\vec{u}_1 + \dots + 0\vec{u}_R$ is clearly in the span as well. Hence the span of these vectors is in fact a subspace

Question 4

Prob. Diabetes = 9.3%

Prob. Inactive genes | Diabetes = 95%

a.) This cannot determined

b.) We must know the probability that someone without diabetes has those same inactive genes

c.) I would be concerned if the inactive genes are very rare amongst the non-Diabetic population.

Question 5.

In this instance, I will model the message delay using an exponential R.V. that is shifted to have minimal delay of t_0

$$\Rightarrow \quad t - t_0 \sim \exp(\theta)$$

$$\Rightarrow \mathbb{P}(t) = \begin{cases} \theta e^{-\theta(t - t_0)} & t \geq t_0 \\ 0 & t < t_0 \end{cases}$$

where t is the R.V. corresponding to the time delay

I use this model because the probability of long delays decays very quickly to capture the rarity of long delays. Beyond this the exponential distribution is one-sided which matches the problem

Question 6.

See code

Question 7.

$$\ell(\theta) = \sum_{i=1}^{N} \left[y_i \log\left(\frac{1}{1 + e^{-\theta^T x}}\right) + (1 - y_i) \log\left(\frac{1}{1 + e^{\theta^T x}}\right) \right]$$

a.)

consider $\left[\nabla \log\left(\frac{1}{1 + e^{-\theta^T x}}\right)\right]_i$

$$= \frac{\partial}{\partial \theta_i}(\theta^T x) \frac{\partial}{\partial u}\left[\log\left(\frac{1}{1 + e^{-u}}\right)\right]_{u = \theta^T x}$$

$$= \frac{x_i}{1 + e^{\theta^T x}}$$

and $\left[\nabla \log\left(\frac{1}{1 + e^{\theta^T x}}\right)\right]_i$

$$= \frac{\partial}{\partial \theta_i}(\theta^T x) \frac{d}{du} \log\left(\frac{1}{1 + e^{u}}\right)\Big|_{u = \theta^T x}$$

$$= \frac{x_i e^{\theta^T x}}{1 + e^{\theta^T x}}$$

$$\Rightarrow \quad \frac{\partial \ell}{\partial x_i} = \sum_{j=1}^{n} \left(y_j \frac{x_i}{1 + e^{\theta^T x}} + (1 - y_j) \frac{x_i e^{\theta^T x}}{1 + e^{\theta^T x}} \right)$$

by linearity.

$$\text{or} \quad \nabla \ell(\theta) = \vec{x} \left[\sum_{j=1}^{n} \left(\frac{1 - e^{\theta^T x}}{1 + e^{\theta^T x}} y_j + \frac{e^{\theta^T x}}{1 + e^{\theta^T x}} \right) \right].$$

b.)

$$\frac{\partial \ell}{\partial x_i \partial x_j} = x_i x_j \sum_{k=1}^{n} \frac{(1 - 2y_k) e^{\theta^T x}}{(1 + e^{\theta^T x})^2}$$

↑ this gives every element of the Hessian. Otherwise it is given by

$$\vec{x}\vec{x}^T \sum_{k=1}^{n} \frac{(1 - 2y_k) e^{\theta^T x}}{(1 + e^{\theta^T x})^2}$$

c.) $\ell(\theta)$ is a scalar

$\nabla \ell(\theta)$ is a vector

H_ℓ is a matrix

↑ Hessian