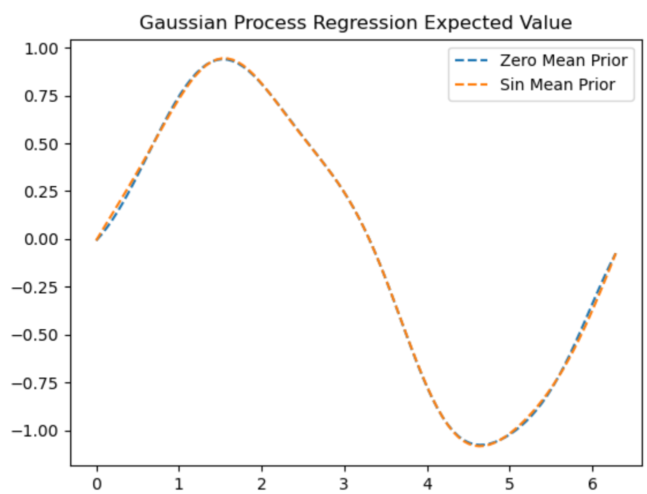
Computational Stastistics HW2

lc3919

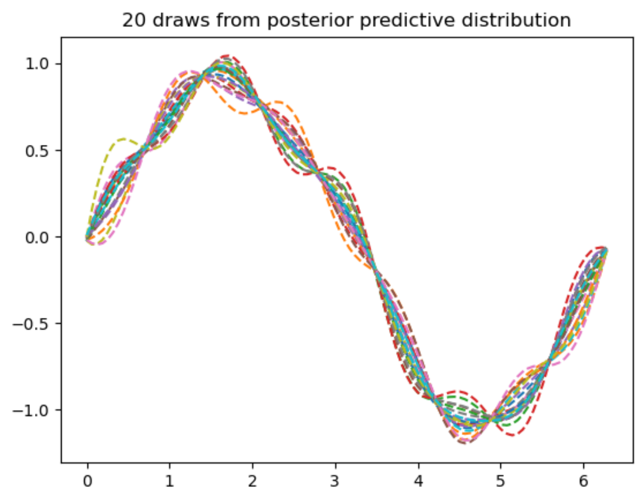
October 2021

Problem 1

a)



b)



Problem 3

We have that $\mathbf{A} \in \mathbb{R}^{m\times n}$ such that $Rank(A) = k < n$. We want to show that the QR factorization of A can be used to find the least squares solution

$$\mathbf{x}^* = \text{argmin}_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$$

We are looking at a tall skinny rank deficient matrix. The QR decomposition of this matrix takes the form,

$$\mathbf{A}\Pi = \mathbf{Q}\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\\ 0 & 0\end{pmatrix}$$

Where Π is some permutation matrix that does not change the norm such that the first k columns are linearly independent. Then $\mathbf{R_1} \in \mathbb{R}^{k\times k}$ and $\mathbf{R_2} \in \mathbb{R}^{(m-k)\times k}$. So we can now write,

$$A = \mathbf{Q}\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\\ 0 & 0\end{pmatrix}\Pi^T$$

Let's now substitute this into our the 2-norm and see what we can do,

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 = \|\mathbf{Q}\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\\ 0 & 0\end{pmatrix}\Pi^T\mathbf{x} - b\|_2$$

$$= \|\begin{pmatrix}\mathbf{R}_1 & \mathbf{R}_2\\ 0 & 0\end{pmatrix}\Pi^T\mathbf{x} - \mathbf{Q}^T b\|_2$$

If we then partition $\Pi^T\mathbf{x} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ and $\mathbf{Q}^T b = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ we can write,

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 = \| \begin{pmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \|$$

$$= \| \begin{pmatrix} \mathbf{R_1} z_1 + \mathbf{R}_2 z_2 - c_1 \\ -c_2 \end{pmatrix} \|$$

$$\|\mathbf{R_1} z_1 + \mathbf{R}_2 z_2 - c_1\|_2 + \|c_2\|_2$$

The first part of this sum is what we can actually control so to minimize $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$,

$$\mathbf{R_1} z_1 + \mathbf{R}_2 z_2 - c_1 = 0$$

$$z_1 = \mathbf{R_1}^{-1}(c_1 - \mathbf{R_2} z_2)$$

And letting $z_2 = 0$ we can find the least squares solution for x,

$$\Pi^T \mathbf{x} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

$$\mathbf{x}^* = \Pi \begin{pmatrix} \mathbf{R_1}^{-1}(c_1 - \mathbf{R_2} z_2) \\ z_2 \end{pmatrix}$$

Problem 4

We have that $\mathbf{A} \in \mathbb{R}^{m \times n}$ where $m < n$ and $Rank(A) = m$. So we have an over-determined system and we want to find the minimum 2-norm least squares solution. For a matrix of the type we have above we can write the singular value decomposition of A as,

$$\mathbf{A} = \underbrace{\begin{pmatrix} \mathbf{U_1} & \mathbf{U_2} \end{pmatrix}}_{\mathbf{U}} \underbrace{\begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} \mathbf{V_1}^T \\ \mathbf{V_2}^T \end{pmatrix}}_{\mathbf{V}^T}$$

Where $\mathbf{U_1} \in \mathbb{R}^{m \times m}, \mathbf{U_2} \in \mathbb{R}^{m \times (n-m)}, \Sigma_1 \in \mathbb{R}^{m \times m}, \mathbf{V_1}^T \in \mathbb{R}^{m \times n}, \mathbf{V_2}^T \in \mathbb{R}^{(n-m) \times n}$ Plugging this into our 2-norm difference we get,

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 = \| \begin{pmatrix} \mathbf{U_1} & \mathbf{U_2} \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{V_1}^T \\ \mathbf{V_2}^T \end{pmatrix} \mathbf{x} - \mathbf{b}\|_2$$

$$= \| \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{V_1}^T \\ \mathbf{V_2}^T \end{pmatrix} \mathbf{x} - \begin{pmatrix} \mathbf{U_1}^T \\ \mathbf{U_2}^T \end{pmatrix} \mathbf{b}\|_2$$

Then letting $\begin{pmatrix} \mathbf{V_1}^T \\ \mathbf{V_2}^T \end{pmatrix} \mathbf{x} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{U_1}^T \\ \mathbf{U_2}^T \end{pmatrix} \mathbf{b} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ we get,

$$\| \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \|_2$$

$$= \| \begin{pmatrix} \Sigma_1 z_1 \\ 0 \end{pmatrix} - \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \|_2$$

$$= \| \begin{pmatrix} \Sigma_1 z_1 - c_1 \\ -c_2 \end{pmatrix} \|_2 = \|\Sigma_1 z_1 - c_1\|_2 + \| - c_2\|_2$$

The first term in the sum which we can control. Setting it to zero we find,

$$z_1 = \Sigma_1^{-1} c_1$$

$$\mathbf{V_1}^T \mathbf{x} = \Sigma_1^{-1} \mathbf{U_1}^T \mathbf{b}$$

$$\mathbf{x} = \mathbf{V}_1 \Sigma_1 \mathbf{U_1}^T \mathbf{b}$$

Which is just the pseudo inverse of A. For any choice of $z_2 \neq 0$ the L-2 norm would no longer be a minimum.