Solution of Homework1

October 11, 2016

problem2

For a given vector $\mathbf{x} \in \mathbb{R}^n$, its norms can be calculated as follows:

$$
\begin{aligned}
\|\mathbf{x}\|_1 &= \sum_{i=1}^{n} |\mathbf{x}_i| \\
\|\mathbf{x}\|_2 &= \sqrt{\sum_{i=1}^{n} \mathbf{x}_i^2} \\
\|\mathbf{x}\|_\infty &= \max_{i=1,\cdots,n} |\mathbf{x}_i|
\end{aligned} \tag{1}
$$

- $\|\mathbf{x}\|_1 \geq \|\mathbf{x}\|_2$:

  We use induction to prove the inequality as follows:

  - When $n = 1$:

    $\|\mathbf{x}\|_1 = |\mathbf{x}_1|$ and $\|\mathbf{x}\|_2 = |\mathbf{x}_1|$. Then, we can see that it is right.

  - When $k = n - 1$ the inequality is right, we need prove the inequality is still correct when $k = n$.

$$
\begin{aligned}
\|\mathbf{x}\|_1^2 &= \left(\sum_{i=1}^{n} |\mathbf{x}_i|\right)^2 \\
&= \left(\sum_{i=1}^{n-1} |\mathbf{x}_i|\right)^2 + 2 \cdot |\mathbf{x}_n| \cdot \sum_{i=1}^{n-1} |\mathbf{x}_i| + |\mathbf{x}_n|^2 \\
&\geq \sum_{i=1}^{n-1} |\mathbf{x}_i|^2 + |\mathbf{x}_n|^2 = \|\mathbf{x}\|_2^2
\end{aligned} \tag{2}
$$

Prove done!

- $\|\mathbf{x}\|_2 \geq \|\mathbf{x}\|_\infty$:

  Firstly, we assume that $\|\mathbf{x}\|_\infty = |\mathbf{x}|_k$.

$$\|\mathbf{x}\|_2^2 = \sum_{i=1}^{n} |\mathbf{x}_i|^2 = \sum_{i=1, i\neq k}^{n} |\mathbf{x}_i|^2 + \mathbf{x}_k^2 \geq \mathbf{x}_k^2 \tag{3}$$

  It's easily find the conclusion is correct.

Therefore, we prove that

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1. \tag{4}$$

problem3

$$\frac{\partial q(\mathbf{x})}{\partial \mathbf{x}} = 2 \cdot \mathbf{A}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) \tag{5}$$

$$\frac{\partial^2 q(\mathbf{x})}{\partial \mathbf{x}^2} = 2\mathbf{A}^T\mathbf{A} \tag{6}$$

The Hessian matrix of $q(\mathbf{x})$ is a positive-semidefinite matrix, so $q(\mathbf{x})$ is convex.

If $\mathbf{B} \in \mathbb{R}^{m\times k}$ and $\mathbf{C} \in \mathbb{R}^{k\times n}$, we have the inequality as follows:

$$(\mathsf{rank}(\mathbf{B}) + \mathsf{rank}(\mathbf{C})) - k \leq \mathsf{rank}(\mathbf{BC}) \leq \mathsf{rank}(\mathbf{B}) + \mathsf{rank}(\mathbf{C}) \tag{7}$$

According to $\mathsf{rank}(\mathbf{A}) = n$, we know that $\mathsf{rank}(\mathbf{A}^T\mathbf{A}) = n$,Then, we can derive its minimizer as follows:

$$\mathbf{x}^* = \left(\mathbf{A}^T\mathbf{A}\right)^{-1} \mathbf{A}^T\mathbf{b}. \tag{8}$$

Since $q(\mathbf{x})$ is convex,it is the global minimizer .

problem4

For sufficiently small stepsize α, we consider the next $\mathbf{x}_\alpha$ can be expressed as follows:

$$\mathbf{x}_\alpha = \mathbf{x} + \alpha \cdot \mathbf{d}. \tag{9}$$

By first-order Taylor series expansion around $\mathbf{x}$, we have

$$\begin{aligned} f(\mathbf{x}_\alpha) &= f(\mathbf{x}) + \alpha \nabla f(\mathbf{x})^T \mathbf{d} + o(\alpha) \\ &\approx f(\mathbf{x}) + \alpha \nabla f(\mathbf{x})^T \mathbf{d} < f(\mathbf{x}) \end{aligned} \tag{10}$$

Prove done!

problem5

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + x_2^2) \end{bmatrix} \tag{11}$$

Then, we calculate for $\nabla f(\mathbf{x})^T \mathbf{p}$.

$$\nabla f(\mathbf{x})^T \mathbf{p} = [2, 0] \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -2 < 0. \tag{12}$$

Therefore, $\mathbf{p}$ is a descent direction. Since $\mathbf{x} + \alpha \mathbf{p} = [1 - \alpha, \alpha]^T$, then $f(\mathbf{x})$ can be represented as

$$f(\alpha) = \left(1 - \alpha + \alpha^2\right)^2. \tag{13}$$

$$\frac{\partial f(\alpha)}{\partial \alpha} = 2(2\alpha - 1)(1 - \alpha + \alpha^2) = 0 \tag{14}$$

We derive three solutions $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{1+\sqrt{3}i}{2}$ and $\alpha_3 = \frac{1-\sqrt{3}i}{2}$. Since we need $\alpha > 0$, then $\alpha^* = \frac{1}{2}$, and $f^* = \frac{9}{16}$.