Summer Reading 2023

Instructor: Greg Ongie - Student: Henrique Medeiros Dos Reis

Summer 2023

Problem 1. Characterize the least squares solution

Solution 1.

$$\begin{aligned} min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{y}\|^2 & \\ min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{y}\|^2 &= (A\mathbf{x} - \mathbf{y})^\top (A\mathbf{x} - \mathbf{y}) \\ &= (\mathbf{x}^\top A^\top - \mathbf{y}^\top)((A\mathbf{x} - \mathbf{y}) \\ &= \mathbf{x}^\top A^\top A\mathbf{x} - \mathbf{x}^\top A^\top \mathbf{y} - \mathbf{y}^\top A\mathbf{x} - \mathbf{y}^\top \mathbf{y} \end{aligned}$$

Now, taking the gradient with respect to $\mathbf{x}$, and setting it equal to 0

$$\nabla \|A\mathbf{x} - \mathbf{y}\|^2 = 2A^\top A\mathbf{x} - 2A^\top \mathbf{y}$$

$$\begin{aligned} 2A^\top A\mathbf{x} - 2A^\top \mathbf{y} = 0 \\ A^\top A\mathbf{x} = A^\top]by \end{aligned}$$

Assuming $A^\top A$ is invertible, $\mathbf{x}_{ls} = (A^\top A)^{-1} A^\top \mathbf{y}$.

Which for image related problems will almost never be the case, so the solution will not be unique. Then we will need some kind of regularization or pseudo inverse.

Problem 2. Prove that if $f''(x) \geq 0$, then f is convex in the sense of $f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$ for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n, \alpha \in [0,1]$

Solution 2.

$$\begin{aligned} f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) &\leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \\ f(\alpha \mathbf{x}_1 + \mathbf{x}_2 - \alpha \mathbf{x}_2) &\leq \alpha f(\mathbf{x}_1) + f(\mathbf{x}_2) - \alpha f(\mathbf{x}_2) \\ f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2) &\leq \alpha (f(\mathbf{x}_1) - f(\mathbf{x}_2)) \\ \frac{f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\alpha} &\leq f(\mathbf{x}_1) - f(\mathbf{x}_2) \end{aligned}$$

As $\alpha \to 0$,

$$\lim_{\alpha \to 0} \frac{f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\alpha}$$

Which is pretty similar to a derivative (can I say gradient here?), but multiplied by $(\mathbf{x}_1 - \mathbf{x}_2)$

$$\Rightarrow \nabla f(\mathbf{x}_2)^\top (\mathbf{x}_1 - \mathbf{x}_2) \leq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$
$$\nabla f(\mathbf{x}_2)^\top (\mathbf{x}_1 - \mathbf{x}_2) + f(\mathbf{x}_2) \leq f(\mathbf{x}_1)$$

Now looking at the Taylor expansion of $f(\mathbf{x}_1)$

$$f(\mathbf{x}_1) = f(\mathbf{x}_2) + \nabla(\mathbf{x}_2)^\top (\mathbf{x}_1 - \mathbf{x}_2) + \frac{1}{2}(\mathbf{x}_1 - \mathbf{x}_2)^\top \nabla^2 f(\mathbf{z})(\mathbf{x}_1 - \mathbf{x}_2) \text{ for some } \mathbf{z} \in [\mathbf{x}_1, \mathbf{x}_2]$$

$$\Rightarrow f(\mathbf{x}_1) - \frac{1}{2}(\mathbf{x}_1 - \mathbf{x}_2)^\top \nabla^2 f(\mathbf{z})(\mathbf{x}_1 - \mathbf{x}_2) = f(\mathbf{x}_2) + \nabla(\mathbf{x}_2)^\top (\mathbf{x}_1 - \mathbf{x}_2)$$

And the term $\frac{1}{2}(\mathbf{x}_1 - \mathbf{x}_2)^\top \nabla^2 f(\mathbf{z})(\mathbf{x}_1 - \mathbf{x}_2)$ is always positive, since $\nabla^2 f(\mathbf{x}) \succeq 0$, then

$$f(\mathbf{x}_1) \geq f(\mathbf{x}_2) + \nabla(\mathbf{x}_2)^\top (\mathbf{x}_1 - \mathbf{x}_2)$$

Which is the exact same equation as we saw above.