Quiz-1 Solutions

Sunday 21st February, 2016

Problem 1.

Problem 2a. f(x_1, x_2) = $1/(x_1x_2)$ on R^2_{++}

To prove f(x_1, x_2) is convex/concave/neither we need to find the hessian first

$$f_{x_1} = \frac{\partial f}{\partial x_1} = -\frac{1}{x_1^2x_2} \qquad f_{x_2} = \frac{\partial f}{\partial x_2} = -\frac{1}{x_1x_2^2}$$

$$f_{x_1x_1} = \frac{\partial f_{x_1}}{\partial x_1} = \frac{2}{x_1^3x_2} \qquad f_{x_2x_2} = \frac{\partial f_{x_2}}{\partial x_2} = \frac{2}{x_1x_2^3}$$

$$f_{x_1x_2} = \frac{\partial f_{x_1}}{\partial x_2} = f_{x_2x_1} = \frac{\partial f_{x_2}}{\partial x_1} = \frac{1}{x_1^2x_2^2}$$

So the Hessian is,

$$\begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{vmatrix} = \begin{vmatrix} \frac{2}{x_1^3x_2} & \frac{1}{x_1^2x_2^2} \\ \frac{1}{x_1^2x_2^2} & \frac{2}{x_1x_2^3} \end{vmatrix}$$

say the eigen values of the hessian are λ_1 and λ_2

as x_1, x_2 on R^2_{++}

$\lambda_1 + \lambda_2$ = trace of the hessian = $(\frac{2}{x_1^3x_2} + \frac{2}{x_1x_2^3}) > 0$ ... (1)

$\lambda_1.\lambda_2$ = The determinant of the hessian = $\frac{3}{x_1^4x_2^4} > 0$ ... (2)

From 1 and 2 we can conclude λ_1 and $\lambda_2 > 0$ and the hessian is positive definite.

Hence, the given function is **convex.**

Problem 2b. Using the property of norm $||a+b||_p \leq ||a||_p + ||b||_p$ (triangle inequality) and $0 \leq \theta \leq 1$,

$$f(x) = ||x||_p$$

$$f(\theta x_1 + (1-\theta)x_2) = ||\theta x_1 + (1-\theta)x_2||_p$$

$$\leq ||\theta x_1||_p + ||(1-\theta)x_2||_p$$

$$\leq \theta||x_1||_p + (1-\theta)||x_2||_p$$

$$\leq \theta f(x_1) + (1-\theta)f(x_2)$$

Hence, the given function is **convex.**

Problem 3.

- $$\mathbf{w}^* = \underset{\mathbf{w}}{\mathbf{argmin}} \|\phi\mathbf{w} - \mathbf{y}\|^{\mathbf{2}} \ \text{ s.t. } \|\mathbf{w}\|_{\mathbf{1}} \le \eta, \tag{1}$$

  where

  $$\|\mathbf{w}\|_1 = \left(\sum_{i=1}^{n} |w_i|\right) \tag{2}$$

- Since $\|\mathbf{w}\|_1$ is not differentiable, one can express (2) as a set of constraints

  $$\sum_{i=1}^{n} \xi_i \le \eta, \ w_i \le \xi_i, \ -w_i \le \xi_i$$

- The resulting problem is a linearly constrained Quadratic optimization problem (LCQP):

  $$\mathbf{w}^* = \underset{\mathbf{w}}{\mathbf{argmin}} \|\phi\mathbf{w} - \mathbf{y}\|^{\mathbf{2}} \ \text{ s.t. } \sum_{\mathbf{i=1}}^{\mathbf{n}} \xi_{\mathbf{i}} \le \eta, \ \mathbf{w_i} \le \xi_{\mathbf{i}}, \ -\mathbf{w_i} \le \xi_{\mathbf{i}} \tag{3}$$

- Lagrangian is

  $$\|\phi\mathbf{w} - \mathbf{y}\|^2 \ + \beta(\sum_{i=1}^{n} \xi_i - \eta) + \sum_{i=1}^{n} \big(\theta_i(\ w_i - \xi_i) + \lambda_i(-w_i - \xi_i)\big)$$

- KKT conditions: Setting gradient wrt $\mathbf{w}$ to $\mathbf{0}$:

  $$2(\phi^T\phi)\mathbf{w} - \mathbf{2}\phi^{\mathbf{T}}\mathbf{y} + (\theta - \lambda) = \mathbf{0}$$

  Setting gradient wrt ξ_i to 0:

  $$\beta - \theta_i - \lambda_i = 0$$

  $$\beta(\sum_{i=1}^{n} \xi_i - \eta) = 0$$

  $$\forall \ i, \ \theta_i(\mathbf{w_i} - \xi_{\mathbf{i}}) = \mathbf{0} \text{ and } \lambda_{\mathbf{i}}(-\mathbf{w_i} - \xi_{\mathbf{i}}) = \mathbf{0}$$

- We have also shown the equivalence of Lasso formulations in (2) and (4):

  $$\mathbf{w}^* = \underset{\mathbf{w}}{\mathbf{argmin}} \|\phi\mathbf{w} - \mathbf{y}\|^{\mathbf{2}} + \lambda\|\mathbf{w}\|_{\mathbf{1}} \tag{4}$$