Coursework

Imperial College London

Department of Computing

477 - Computational Optimisation

Author:
Jiahao Lin (CID: 00837321)

Date: November 6, 2016

1 Part 1

1.1 1

1) To prove is $\log\sum_{k=1}^{10} exp(B_{jk})$ is convex, we define:

$$f(B_j) = \log\sum_{k=1}^{10} exp(B_{jk}) \tag{1}$$

$$f(x) = \log\sum_{k=1}^{10} exp(x_k) \tag{2}$$

First we need to compute the Hessian of $f(x)$:

$$\nabla f(x) = \frac{1}{1^\top Z} \times z \tag{3}$$

$$\nabla^2 f(x) = \frac{1}{1^\top Z} \times diag(Z) - \frac{Z^\top Z}{(1^\top Z)^2} \tag{4}$$

$$where \quad Z = \sum_{k=1}^{10} exp(x_k) \tag{5}$$

For convexity we need to prove the Hessian is positive semi-definite, we need to prove:

$$\nabla^2 f(x) \geq 0 \tag{6}$$

$$v^\top \nabla^2 f(x) v >= 0 \tag{7}$$

$$v^\top \left(\frac{1}{1^\top Z} \times diag(Z) - \frac{Z^\top Z}{(1^\top Z)^2} \right) v >= 0 \tag{8}$$

$$\frac{(\sum_{k=1}^{10} Z_k V_k^2)\sum_{k=1}^{10} Z_k - \sum_{k=1}^{10}(Z_k v_k)^2}{(1^\top Z)^2} >= 0 \tag{9}$$

According to Cauchy-Schwartz inequality:

$$\sum_{k=1}^{10}(Z_k v_k)^2 \leq \sum_{k=1}^{10} Z_k \times \sum_{k=1}^{10}(Z_k V_k^2) \tag{10}$$

There for the Hessian is positive semi-definite holds, the Log-Sum-Exp function is convex.

2) Denote affine mapping on β_k as $g(\beta_k)$:

$$g(\beta_k) = x_i^\top \beta_k \tag{11}$$

And denote Log-Sum-Exp function as $f(x_k)$:

$$f(x_k) = \log\sum_{k=1}^{10} exp(x_k) \tag{12}$$

Now we know that $f(x_k)$ is convex, and we want to prove that $f \circ g$ is convex as well, from the definition of convexity, we need to prove:

$$f \circ g(\alpha x + (1-\alpha)y) \leq \alpha f \circ g(x) + (1-\alpha) f \circ g(y) \tag{13}$$
$$for \quad any \quad x, \quad y \quad and \quad \alpha \in [0,1] \tag{14}$$

Since we know:

$$g(\alpha x + (1-\alpha)y) = \alpha g(x) + (1-\alpha) g(y) \tag{15}$$

We can prove that:

$$f \circ g(\alpha x + (1-\alpha)y) = f(g(\alpha x + (1-\alpha)y)) \tag{16}$$
$$= f(\alpha g(x) + (1-\alpha) g(y)) \tag{17}$$
$$\leq \alpha f \circ g(x) + (1-\alpha) f \circ g(y) \tag{18}$$

Now we know that $\log \sum_{k=1}^{10} exp(x_i^\top \beta_k)$ is convex as well.

3) Define function $f(\beta_y)$:

$$f(\beta_{y_i}) = -x_i^\top \beta_{y_i+1} \tag{19}$$

We can show that:

$$f(\alpha \beta_{x_i+1} + (1-\alpha)\beta_{y_i+1}) = -x_i^\top (\alpha \beta_{x_i+1} + (1-\alpha)\beta_{y_i} + 1) \tag{20}$$
$$= -\alpha x_i^\top \beta_{x_i+1} - (1-\alpha) x_i^\top \beta_{y_i+1} \tag{21}$$
$$= \alpha f(\beta_{x_i+1}) + (1-\alpha) f(\beta_{y_i+1}) \tag{22}$$

Therefore $f(\beta_y)$ is convex.

4) We can define a function for $\ell 1$ Regularisation $||\beta_k||_1$:

$$||\beta_k||_1 = f(\beta_k) = \sum_{i=1}^{N} |\beta_{ki}| \tag{23}$$

To prove the convexity of $f(\beta_k)$, we can show that:

$$f(\alpha \beta_k + (1-\alpha)\beta_j) = \sum_{i=1}^{N} |\alpha \beta_k + (1-\alpha)\beta_j| \tag{24}$$
$$= \alpha \sum_{i=1}^{N} |\beta_{ki}| + (1-\alpha) \sum_{i=1}^{N} |\beta_{ji}| \tag{25}$$
$$= \alpha f(\beta_k) + (1-\alpha) f(\beta_j) \tag{26}$$

Therefore $f(\beta_k)$ is convex.

5)