

COURSEWORK

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

477 - Computational Optimisation

Author:

Jiahao Lin (CID: 00837321)

Date: November 6, 2016

1 Part 1

1.1 Q.1

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

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

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

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

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

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

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

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

$$\nabla^2 f(x) \geq 0 \quad (6)$$

$$v^\top \nabla^2 f(x) v \geq 0 \quad (7)$$

$$v^\top \left(\frac{1}{1^\top Z} \times \text{diag}(Z) - \frac{Z^\top Z}{(1^\top Z)^2} \right) v \geq 0 \quad (8)$$

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

According to Cauchy-Schwartz inequality:

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

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

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

$$g(\beta_k) = x_i^\top \beta_k \quad (11)$$

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

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

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:

$$f \circ g(\alpha x + (1 - \alpha)y) \leq \alpha f \circ g(x) + (1 - \alpha)f \circ g(y) \quad (13)$$

$$\text{for any } x, y \text{ and } \alpha \in [0, 1] \quad (14)$$

Since we know:

$$g(\alpha x + (1 - \alpha)y) = \alpha g(x) + (1 - \alpha)g(y) \quad (15)$$

We can prove that:

$$f \circ g(\alpha x + (1 - \alpha)y) = f(g(\alpha x + (1 - \alpha)y)) \quad (16)$$

$$= f(\alpha g(x) + (1 - \alpha)g(y)) \quad (17)$$

$$\leq \alpha f \circ g(x) + (1 - \alpha)f \circ g(y) \quad (18)$$

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

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

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

We can show that:

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

$$= -\alpha x_i^\top \beta_{x_i+1} - (1 - \alpha)x_i^\top \beta_{y_i+1} \quad (21)$$

$$= \alpha f(\beta_{x_i+1}) + (1 - \alpha)f(\beta_{y_i+1}) \quad (22)$$

Therefore $f(\beta_y)$ is convex.

4) We can define a function for ℓ_1 Regularisation $\|\beta_k\|_1$:

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

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

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

$$= \alpha \sum_{i=1}^N |\beta_{ki}| + (1 - \alpha) \sum_{i=1}^N |\beta_{ji}| \quad (25)$$

$$= \alpha f(\beta_k) + (1 - \alpha)f(\beta_j) \quad (26)$$

Therefore $f(\beta_k)$ is convex.

5) To prove the following function is convex:

$$\sum_{i=1}^m \left(\log \sum_{k=1}^{10} \exp(x_i^\top \beta_k) - x_i^\top \beta_{y_i+1} \right) + \lambda \sum_{k=1}^{10} \|\beta_k\|_1 \quad (27)$$

We need to prove by decomposition, in Q2, we have shown the following term to be convex:

$$\log \sum_{k=1}^{10} \exp(x_i^\top \beta_k) \quad (28)$$

and in Q3, we have shown the following term is convex:

$$-x_i^\top \beta_{y_i+1} \quad (29)$$

Therefore according to sum of convex function is also convex, we can get the sum of both is convex as well:

$$\log \sum_{k=1}^{10} \exp(x_i^\top \beta_k) - x_i^\top \beta_{y_i+1} \quad (30)$$

Following the same lemma, the summation term is also convex:

$$\sum_{i=1}^m \left(\log \sum_{k=1}^{10} \exp(x_i^\top \beta_k) - x_i^\top \beta_{y_i+1} \right) \quad (31)$$

In Q4, we have shown the ℓ_1 Regularisation term is convex:

$$\|\beta_k\|_1 \quad (32)$$

Then by sum of convex function is convex and affine mapping of convex function is convex, we get that the second term of optimisation problem function is convex:

$$\lambda \sum_{k=1}^{10} \|\beta_k\|_1 \quad (33)$$

Combining the convex first term and convex second term, the whole optimisation problem is convex:

$$\sum_{i=1}^m \left(\log \sum_{k=1}^{10} \exp(x_i^\top \beta_k) - x_i^\top \beta_{y_i+1} \right) + \lambda \sum_{k=1}^{10} \|\beta_k\|_1 \quad (34)$$

2 Part 2

2.1 Q.2

1)

$$h(B) = \lambda \sum_{k=1}^{10} \|\beta_k\|_1 \quad (35)$$

$$= \lambda \sum_{k=1}^{10} \sum_{i=1}^n |\beta_{ik}| \quad (36)$$

Since absolute value is not differentiable because it has a kink at 0, the whole $h(\mathcal{B})$ function is not differentiable. So $h(\mathcal{B}) \notin \mathcal{C}^1$.

2)

3)

4)

5)

6)

7)