Homework assignments Week12 R10946017 許鴻彥

③ Derivation of the Lipschitz constant....

The correct answer is (c).

Sol: gradient algorithm: $\beta^{r+1} = \beta^{r} - c\nabla\mathcal{L}(\beta^{r})$

Let $c = \frac{1}{L}$, where L is the Lipschitz constant.

If $c = \frac{1}{L}$ is small enough, then the G.D. process will make $\mathcal{L}(\beta)$ decrease.

Assume $\mathcal{L}(\beta)$ is Lipschitz continuous, then

$$\|\nabla\mathcal{L}(\beta^{r}) - \nabla\mathcal{L}(\beta^{r+1})\| \leq L\|\beta^{r} - \beta^{r+1}\|$$

As $\beta^{r} = \beta^{r+1}$, the Lipschitz continuity of the gradient:

$$\nabla^2\mathcal{L}(\beta^{r}) = \underbrace{H(\beta^{r})}_{\text{hessian matrix}} \leq LI$$

Therefore, the eigenvalues of the Hessian matrix are bounded above by L, and the minimum L is the maximum eigenvalue of the gram matrix.