Homework 4

Clint Guo

2018-12-04

Contents

1. CASL page 143, question 2

We can generate a very simple example where the linear Hessian (X^tX) is well-conditioned but its logistic variation is not (with the generation of corresponding β and p).

Let's see a toy linear model example with $n = p = 2$. We set X as

$$X = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

The linear Hessian (X^tX) is very easy to computed given X is a diagnal matrix.

$$H_1 = X^tX = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$$

The condition number of a matrix is the ratio of its largest and smallest singular value.

$$cond(H_1) = \frac{\sqrt{4}}{\sqrt{1}} = 2$$

While the Hessian is well-conditioned, we set $\beta = (0, -100)^t$ such that

$$y = X * \beta = \begin{pmatrix} 0 \\ -100 \end{pmatrix}$$

$$p_1 = \frac{e^0}{1 + e^0} = 0.5$$

$$p_2 = \frac{e^{-100}}{1 + e^{-100}} = 3.7 * 10^{-44}$$

We set $p_1 = 0.5$ is due to the fact that the expression $p(1 - p)$ reaches its maximum 0.25 at $p = 0.5$.

The logistic Hessian is computed as follow.

$$\begin{aligned} H_2 &= X^t * diag(p * (1 - p)) * X \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} * \begin{pmatrix} 0.5 * (1 - 0.5) & 0 \\ 0 & 3.7 * 10^{-44} * (1 - 3.7 * 10^{-44}) \end{pmatrix} * \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 3.7 * 10^{-44} \end{pmatrix} \end{aligned}$$

The condition number is then computed.

$$cond(H_2) = \frac{\sqrt{1}}{\sqrt{3.7 * 10^{-44}}} = 5.2 * 10^{21}$$