Determinantal Regularization for Ensemble Variable Selection

Supplementary Material

In the following theorem, we derive the form of the Hessian matrix $\mathbf{H}_k$ which first appears in equation (8) of the main paper.

Theorem 1. *The Hessian* $\mathbf{H}_k$ *of* $\log \pi(\mathbf{Y}, \boldsymbol{\beta})$ *has the following form*

$$\mathbf{H}_k = -[\mathbf{X}'\mathbf{X} + \mathbf{D}^*(\boldsymbol{\beta})] - 2\left(\frac{1}{v_1} - \frac{1}{v_0}\right) \operatorname{diag}\left\{\beta_i \frac{\partial p^*(\beta_i)}{\partial \beta_i}\right\}_{i=1}^p - \frac{1}{2}\left(\frac{1}{v_1} - \frac{1}{v_0}\right) \operatorname{diag}\left\{\beta_i^2 \frac{\partial^2 p^*(\beta_i)}{\partial \beta_i^2}\right\}_{i=1}^p . \tag{0.1}$$

where $\mathbf{D}^*(\boldsymbol{\beta}) = \operatorname{diag}\{d^*(\beta_i)\}_{i=1}^p$ *with*

$$d^*(\beta_i) = \frac{1}{v_1} p^*(\beta_i) + \frac{1}{v_0}[1 - p^*(\beta_i)]$$

and $p^*(\beta_i) = \mathsf{P}(\gamma_i = 1 | \beta_i, \theta)$.

Proof. To begin, we have

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\beta}} \log \pi(\boldsymbol{Y}, \boldsymbol{\beta}) &= \frac{\partial}{\partial \boldsymbol{\beta}} \log\left(\sum_{\boldsymbol{\gamma}} \pi(\boldsymbol{Y}, \boldsymbol{\beta} \mid \boldsymbol{\gamma}) \pi(\boldsymbol{\gamma})\right) \\
&= \frac{1}{\pi(\boldsymbol{Y}, \boldsymbol{\beta})} \sum_{\boldsymbol{\gamma}} \frac{\partial}{\partial \boldsymbol{\beta}} \pi(\boldsymbol{Y}, \boldsymbol{\beta} \mid \boldsymbol{\gamma}) \pi(\boldsymbol{\gamma}) = \sum_{\boldsymbol{\gamma}} \frac{\partial}{\partial \boldsymbol{\beta}} \log \pi(\boldsymbol{Y}, \boldsymbol{\beta} \mid \boldsymbol{\gamma}) \pi(\boldsymbol{\gamma} \mid \boldsymbol{Y}, \boldsymbol{\beta}) \\
&= \frac{\partial}{\partial \boldsymbol{\beta}} \left\{-\frac{1}{2}\left(\mathbf{Y}'\mathbf{Y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}\right) - \frac{1}{2}\boldsymbol{\beta}'\mathbf{D}(\boldsymbol{\beta})\boldsymbol{\beta}\right\} \\
&= \mathbf{XY} - [\mathbf{X}'\mathbf{X} + \mathbf{D}(\boldsymbol{\beta})]\boldsymbol{\beta} - \frac{1}{2}\left(\frac{1}{v_1} - \frac{1}{v_0}\right) \operatorname{vec}\left\{\beta_i^2 \frac{\partial p^*(\beta_i)}{\partial \beta_i}\right\}_{i=1}^p .
\end{aligned}$$

$$\frac{\partial^2}{\partial \boldsymbol{\beta} \boldsymbol{\beta}'} \log \pi(\mathbf{Y}, \boldsymbol{\beta}) = -[\mathbf{X}'\mathbf{X} + \mathbf{D}(\boldsymbol{\beta})] - 2\left(\frac{1}{v_1} - \frac{1}{v_0}\right) \operatorname{diag}\left\{\beta_i \frac{\partial p^*(\beta_i)}{\partial \beta_i}\right\}_{i=1}^p \tag{0.2}$$

$$- \frac{1}{2}\left(\frac{1}{v_1} - \frac{1}{v_0}\right) \operatorname{diag}\left\{\beta_i^2 \frac{\partial^2 p^*(\beta_i)}{\partial \beta_i^2}\right\}_{i=1}^p \tag{0.3}$$