Semiparametric Mixed Effects Probit Regression for Binary Data

Brian M. Brost

08 March 2016

Description

A semiparametric regression model with mixed effects for binary data using the probit link.

Implementation

The file probit.semipar.varying.coef.sim.R simulates data according to the model statement presented below, and probit.semipar.varying.coef.mcmc.R contains the MCMC algorithm for model fitting.

Model statement

Let $y_j(t)$ denote longitudinal data observed at times $t \in \mathcal{T}$ in group j, for $j = 1, \ldots, J$. Note that observations are binary; therefore $y_j(t) \in \{0, 1\}$. Let n_j denote the umber of observations in group j. Furthermore, let $\mathbf{x}_j(t)$ be a vector of qX covariates (including the intercept) that are associated with $y_j(t)$, and $\boldsymbol{\beta}_j$ be the corresponding vector of group-level coefficients (i.e., they vary with j but share a common population-level distribution) for which inference is desired. Let $\mathbf{w}_j(t)$ be a vector of qW basis functions evaluated at time $t \in \mathcal{T}$ and $\boldsymbol{\alpha}_j$ be the vector of basis coefficients for group j. Note that the linear combination $\mathbf{w}_j(t)\boldsymbol{\alpha}_j$ models non-linear patterns or dependence non-parametrically.

$$
\begin{aligned}
y_j(t) &\sim \begin{cases} 0, & v_j(t) \leq 0 \\ 1, & v_j(t) > 1 \end{cases} \\
v_j(t) &\sim \mathcal{N}\left(\mathbf{x}_j'(t)\boldsymbol{\beta}_j + \mathbf{w}_j(t)\boldsymbol{\alpha}_j, \mathbf{1}\right) \\
\boldsymbol{\beta}_j &\sim \mathcal{N}\left(\boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}\right) \\
\boldsymbol{\mu}_\beta &\sim \mathcal{N}\left(\mathbf{0}, \sigma_\beta^2\mathbf{I}\right) \\
\boldsymbol{\alpha}_j &\sim \mathcal{N}(\mathbf{0}, \sigma_{\alpha_j}^2\mathbf{I}) \\
\boldsymbol{\Lambda}^{-1} &\sim \text{Wish}\left(\mathbf{S}_0^{-1}, \nu\right) \\
\sigma_{\alpha_j}^2 &\sim \text{IG}(r, q)
\end{aligned}
$$

Models of this type are typically fit using a large number of basis vectors, more than necessary to approximate non-linear trends or dependence; therefore, regularization (e.g., a ridge penalty) is necessary. The parameter $\sigma_{\alpha_j}^2$ can be selected using cross-validation or some model selection criterion (e.g., DIC); alternatively, we model the variance of the basis coefficients ($\sigma_{\alpha_j}^2$) in order to shrink $\boldsymbol{\alpha}_j$ toward 0 where appropriate.

Full conditional distributions

Observation model auxiliary variable ($v_j(t)$):

$$
\begin{aligned}
[v_j(t) \mid \cdot] &\propto [y_j(t) \mid v_j(t)]\left[v_j(t) \mid \mathbf{x}_j'(t)\boldsymbol{\beta}_j + \mathbf{w}_j(t)\boldsymbol{\alpha}_j, \mathbf{1}\right] \\
&\propto \left(1_{\{y_j(t)=0\}}1_{\{v_j(t)\leq 0\}} + 1_{\{y_j(t)=1\}}1_{\{v_j(t)>0\}}\right) \times \mathcal{N}\left(v_j(t) \mid \mathbf{x}_j'(t)\boldsymbol{\beta}_j + \mathbf{w}_j(t)\boldsymbol{\alpha}_j, \mathbf{1}\right) \\
&= \begin{cases} \mathcal{TN}\left(\mathbf{x}_j'(t)\boldsymbol{\beta}_j + \mathbf{w}_j(t)\boldsymbol{\alpha}_j, \mathbf{1}\right)_{-\infty}^{0}, & y_j(t) = 0 \\ \mathcal{TN}\left(\mathbf{x}_j'(t)\boldsymbol{\beta}_j + \mathbf{w}_j(t)\boldsymbol{\alpha}_j, \mathbf{1}\right)_{0}^{\infty}, & y_j(t) = 1 \end{cases}
\end{aligned}
$$

Regression coefficients ($\boldsymbol{\beta}_j$):

$$\begin{aligned}
\left[\boldsymbol{\beta}_j \mid \cdot\right] &\propto \left[\mathbf{v}_j \mid \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j, \mathbf{1}\right]\left[\boldsymbol{\beta}_j \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}\right] \\
&\propto \mathcal{N}\left(\mathbf{v}_j \mid \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j, , \mathbf{1}\right)\mathcal{N}\left(\boldsymbol{\beta}_j \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}\right) \\
&\propto \exp\left\{-\frac{1}{2}\left(\mathbf{v}_j - \left(\mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j\right)\right)'\left(\mathbf{v}_j - \left(\mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j\right)\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)'\boldsymbol{\Lambda}^{-1}\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\left(\mathbf{v}_j - \mathbf{W}_j\boldsymbol{\alpha}_j\right) - \mathbf{X}_j\boldsymbol{\beta}_j\right)'\left(\left(\mathbf{v}_j - \mathbf{W}_j\boldsymbol{\alpha}_j\right) - \mathbf{X}_j\boldsymbol{\beta}_j\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)'\boldsymbol{\Lambda}^{-1}\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\left(\mathbf{v}_j - \mathbf{W}_j\boldsymbol{\alpha}_j\right)'\mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\beta}_j'\mathbf{X}_j'\mathbf{X}_j\boldsymbol{\beta}_j\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(-2\left(\boldsymbol{\mu}_\beta'\boldsymbol{\Lambda}^{-1}\right)\boldsymbol{\beta}_j + \boldsymbol{\beta}_j'\boldsymbol{\Lambda}^{-1}\boldsymbol{\beta}_j\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\left(\left(\mathbf{v}_j - \mathbf{W}_j\boldsymbol{\alpha}_j\right)'\mathbf{X}_j + \boldsymbol{\mu}_\beta'\boldsymbol{\Lambda}^{-1}\right)\boldsymbol{\beta}_j + \boldsymbol{\beta}_j'\left(\mathbf{X}_j'\mathbf{X}_j + \boldsymbol{\Lambda}^{-1}\right)\boldsymbol{\beta}_j\right)\right\} \\
&= \mathcal{N}(\mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1}),
\end{aligned}$$

where $\mathbf{A} = \mathbf{X}_j{}'\mathbf{X}_j + \boldsymbol{\Lambda}^{-1}$, $\mathbf{b}' = \left(\mathbf{v}_j - \mathbf{W}_j\boldsymbol{\alpha}_j\right)'\mathbf{X}_j + \boldsymbol{\mu}_\beta'\boldsymbol{\Lambda}^{-1}$, $\mathbf{v}_j' = \{v_j(t) : t \in \mathcal{T}\}$ (i.e., the vector collecting $v_j(t)$ for all times $t \in \mathcal{T}$), $\mathbf{X}_j$ is an $n_j \times qX$ matrix collecting the vectors $\mathbf{x}_j(t)$ for all times $t \in \mathcal{T}$, and similarly $\mathbf{W}_j$ is an $n_j \times qW$ basis expansion collecting the vectors $\mathbf{w}_j(t)$ for all times $t \in \mathcal{T}$.

Basis coefficients ($\boldsymbol{\alpha}_j$):

$$\begin{aligned}
\left[\boldsymbol{\alpha}_j \mid \cdot\right] &\propto \left[\mathbf{v}_j \mid \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j, \mathbf{1}\right]\left[\boldsymbol{\alpha}_j \mid \mathbf{0}, \sigma^2_{\alpha_j}\right] \\
&\propto \mathcal{N}\left(\mathbf{v}_j \mid \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j, \mathbf{1}\right)\mathcal{N}\left(\boldsymbol{\alpha}_j \mid \mathbf{0}, \sigma^2_{\alpha_j}\right) \\
&\propto \exp\left\{-\frac{1}{2}\left(\mathbf{v}_j - \left(\mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j\right)\right)'\left(\mathbf{v}_j - \left(\mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{W}_j\boldsymbol{\alpha}_j\right)\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\alpha}_j - \mathbf{0}\right)'\left(\sigma^2_{\alpha_j}\mathbf{I}\right)^{-1}\left(\boldsymbol{\alpha}_j - \mathbf{0}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\left(\mathbf{v}_j - \mathbf{X}_j\boldsymbol{\beta}_j\right) - \mathbf{W}_j\boldsymbol{\alpha}_j\right)'\left(\left(\mathbf{v}_j - \mathbf{X}_j\boldsymbol{\beta}_j\right) - \mathbf{W}_j\boldsymbol{\alpha}_j\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\alpha}_j - \mathbf{0}\right)'\left(\sigma^2_{\alpha_j}\mathbf{I}\right)^{-1}\left(\boldsymbol{\alpha}_j - \mathbf{0}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\left(\mathbf{v}_j - \mathbf{X}_j\boldsymbol{\beta}_j\right)'\mathbf{W}_j\boldsymbol{\alpha}_j + \boldsymbol{\alpha}_j'\mathbf{W}_j'\mathbf{W}_j\boldsymbol{\alpha}_j\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\alpha}_j'\left(\sigma^2_{\alpha_j}\mathbf{I}\right)^{-1}\boldsymbol{\alpha}_j\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\left(\left(\mathbf{v}_j - \mathbf{X}_j\boldsymbol{\beta}_j\right)'\mathbf{W}_j\right)\boldsymbol{\alpha}_j + \boldsymbol{\alpha}_j'\left(\mathbf{W}_j'\mathbf{W}_j + \left(\sigma^2_{\alpha_j}\mathbf{I}\right)^{-1}\right)\boldsymbol{\alpha}_j\right)\right\} \\
&= \mathcal{N}(\mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1}),
\end{aligned}$$

where $\mathbf{A} = \mathbf{W}_j'\mathbf{W}_j + \left(\sigma^2_{\alpha_j}\mathbf{I}\right)^{-1}$ and $\mathbf{b}' = \left(\mathbf{v}_j - \mathbf{X}_j\boldsymbol{\beta}_j\right)'\mathbf{W}_j$, $\mathbf{v}_j' = \{v_j(t) : t \in \mathcal{T}\}$, $\mathbf{X}_j$ is an $n_j \times qX$ matrix collecting the vectors $\mathbf{x}_j(t)$ for all times $t \in \mathcal{T}$, and similarly $\mathbf{W}_j$ is an $n_j \times qW$ basis expansion collecting the vectors $\mathbf{w}_j(t)$ for all times $t \in \mathcal{T}$.

Mean of regression coefficients ($\boldsymbol{\mu}_\beta$):

$$
\begin{aligned}
[\boldsymbol{\mu}_\beta \mid \cdot] &\propto \prod_{j=1}^{J} [\boldsymbol{\beta}_j \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}] [\boldsymbol{\mu}_\beta \mid \mathbf{0}, \sigma^2_\beta] \\
&\propto \prod_{j=1}^{J} \mathcal{N}(\boldsymbol{\beta}_j \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}) \mathcal{N}(\boldsymbol{\mu}_\beta \mid \mathbf{0}, \sigma^2_\beta \mathbf{I}) \\
&\propto \exp\left\{\sum_{j=1}^{J}\left(-\frac{1}{2}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)' \boldsymbol{\Lambda}^{-1} (\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}(\boldsymbol{\mu}_\beta - \mathbf{0})' (\sigma^2_\beta\mathbf{I})^{-1} (\boldsymbol{\mu}_\beta - \mathbf{0})\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\left(\sum_{j=1}^{J}\boldsymbol{\beta}_j'\boldsymbol{\Lambda}^{-1}\right)\boldsymbol{\mu}_\beta + \boldsymbol{\mu}_\beta'(J\boldsymbol{\Lambda}^{-1})\boldsymbol{\mu}_\beta\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\mu}_\beta'(\sigma^2_\beta\mathbf{I})^{-1}\boldsymbol{\mu}_\beta\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\left(\sum_{j=1}^{J}\boldsymbol{\beta}_j'\boldsymbol{\Lambda}^{-1}\right)\boldsymbol{\mu}_\beta + \boldsymbol{\mu}_\beta'\left(J\boldsymbol{\Lambda}^{-1} + (\sigma^2_\beta\mathbf{I})^{-1}\right)\boldsymbol{\mu}_\beta\right)\right\} \\
&= \mathcal{N}(\mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1}),
\end{aligned}
$$

where $\mathbf{A} = J\boldsymbol{\Lambda}^{-1} + \left(\sigma^2_\beta\mathbf{I}\right)^{-1}$ and $\mathbf{b}' = \boldsymbol{\beta}'\boldsymbol{\Lambda}^{-1}$, where $\boldsymbol{\beta}$ is the vector sum $\sum_{j=1}^{J}\boldsymbol{\beta}_j$.

Variance-covariance of regression coefficients ($\boldsymbol{\Lambda}$):

$$
\begin{aligned}
[\boldsymbol{\Lambda} \mid \cdot] &\propto \prod_{j=1}^{J} \left[\boldsymbol{\beta}_j \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}\right] \left[\boldsymbol{\Lambda} \mid \mathbf{S}_0, \nu\right] \\
&\propto \prod_{j=1}^{J} \mathcal{N}\left(\boldsymbol{\beta}_j \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Lambda}\right) \text{Wish}\left(\boldsymbol{\Lambda} \mid \mathbf{S}_0, \nu\right) \\
&\propto |\boldsymbol{\Lambda}|^{-\frac{J}{2}} \exp\left\{-\frac{1}{2}\sum_{j=1}^{J}\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)' \boldsymbol{\Lambda}^{-1}\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\right\} \\
&\qquad \times |\mathbf{S}_0|^{-\frac{\nu}{2}} |\boldsymbol{\Lambda}|^{-\frac{\nu - qX - 1}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left(\mathbf{S}_0 \boldsymbol{\Lambda}^{-1}\right)\right\} \\
&\propto |\boldsymbol{\Lambda}|^{-\frac{J+\nu-qX-1}{2}} \exp\left\{-\frac{1}{2}\left[\sum_{j=1}^{J}\text{tr}\left(\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)' \boldsymbol{\Lambda}^{-1}\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\right) + \text{tr}\left(\mathbf{S}_0 \boldsymbol{\Lambda}^{-1}\right)\right]\right\} \\
&\propto |\boldsymbol{\Lambda}|^{-\frac{J+\nu-qX-1}{2}} \exp\left\{-\frac{1}{2}\left[\sum_{j=1}^{J}\text{tr}\left(\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)' \boldsymbol{\Lambda}^{-1}\right) + \text{tr}\left(\mathbf{S}_0 \boldsymbol{\Lambda}^{-1}\right)\right]\right\} \\
&\propto |\boldsymbol{\Lambda}|^{-\frac{J+\nu-qX-1}{2}} \exp\left\{-\frac{1}{2}\left[\text{tr}\left(\sum_{j=1}^{J}\left(\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)'\right) \boldsymbol{\Lambda}^{-1} + \mathbf{S}_0 \boldsymbol{\Lambda}^{-1}\right)\right]\right\} \\
&\propto |\boldsymbol{\Lambda}|^{-\frac{J+\nu-qX-1}{2}} \exp\left\{-\frac{1}{2}\left[\text{tr}\left(\sum_{j=1}^{J}\left(\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)'\right) + \mathbf{S}_0\right) \boldsymbol{\Lambda}^{-1}\right]\right\} \\
&= \text{Wish}\left(\left(\sum_{j=1}^{J}\left(\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)\left(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta\right)'\right) + \mathbf{S}_0\right)^{-1}, J+\nu\right).
\end{aligned}
$$

Variation in basis coefficients ($\sigma^2_{\alpha_j}$):

$$
\begin{aligned}
\left[\sigma^2_{\alpha_j} \mid \cdot\right] &\propto \left[\boldsymbol{\alpha}_j \mid \mathbf{0}, \sigma^2_{\alpha_j}\right]\left[\sigma^2_{\alpha_j}\right] \\
&\propto \mathcal{N}\left(\boldsymbol{\alpha}_j \mid \mathbf{0}, \sigma^2_{\alpha_j}\mathbf{I}\right) \text{IG}\left(r, q\right) \\
&\propto \left|\sigma^2_{\alpha_j}\mathbf{I}\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(\left(\boldsymbol{\alpha}_j - \mathbf{0}\right)'\left(\sigma^2_{\alpha_j}\mathbf{I}\right)^{-1}\left(\boldsymbol{\alpha}_j - \mathbf{0}\right)\right)\right\} \times \\
&\qquad \left(\sigma^2_{\alpha_j}\right)^{-(q+1)} \exp\left\{-\frac{1}{r\sigma^2_{\alpha_j}}\right\} \\
&\propto \left(\sigma^2_{\alpha_j}\right)^{-qW/2} \exp\left\{-\frac{1}{2\sigma^2_{\alpha_j}}\boldsymbol{\alpha}_j'\boldsymbol{\alpha}_j\right\} \\
&\qquad \left(\sigma^2_{\alpha_j}\right)^{-(q+1)} \exp\left\{-\frac{1}{r\sigma^2_{\alpha_j}}\right\} \\
&\propto \left(\sigma^2_{\alpha}\right)^{-(qW/2+q+1)} \exp\left\{-\frac{1}{\sigma^2_{\alpha_j}}\left(\frac{\boldsymbol{\alpha}_j'\boldsymbol{\alpha}_j}{2} + \frac{1}{r}\right)\right\} \\
&= \text{IG}\left(\left(\frac{\boldsymbol{\alpha}_j{}'\boldsymbol{\alpha}_j}{2} + \frac{1}{r}\right)^{-1}, \frac{qW}{2} + q\right)
\end{aligned}
$$

where qW is the number of columns in $\mathbf{W}$.