First Derivatives of the log-L for the multivariate probit model

Vargas Barrenechea, Martin

13 October 2007

First Derivatives of the log-L for the multivariate probit model

Martin H. Vargas Barrenechea *

October 13, 2007

Abstract

In this work we find first derivatives for the log likelihood function of the multivariate probit model.

1 Introduction

The natural extension of the univariate probit model is the multivariate probit model (MVPM) that consists of a system of simultaneous equations of several non-observable dependent variables, in the case of the L-variate probit model the structure is the following:

$$\mathbf{y_i^*} = \begin{pmatrix} y^*_{i,1} \\ y^*_{i,2} \\ \vdots \\ y^*_{i,L} \end{pmatrix} = \begin{pmatrix} x_{i,1}\beta + \varepsilon_{i,1} \\ x_{i,2}\beta + \varepsilon_{i,2} \\ \vdots \\ x_{i,L}\beta + \varepsilon_{i,L} \end{pmatrix} \quad y_{i,l} = \begin{cases} 1 & \text{if } y^*_{i,l} > 0 \\ 0 & \text{if } y^*_{i,l} \le 0 \end{cases}$$

where $\mathbf{y_i^*}$ is a $L \times 1$ vector of non-observable variables, $x_{i,l}$ is a vector $1 \times k_l$ of characteristics of the individual/observation i at the equation l, β_l is a coefficient vector $k_l \times 1$ and $\varepsilon_{i,l}$ is an error.

By stacking the errors $\varepsilon_{i,l}$ we define $\varepsilon_{\mathbf{i}} = (\varepsilon_{i,1}, \ldots, \varepsilon_{i,L})' \sim N(0, P)$ where P is a symmetric matrix $L \times L$ of pairwise correlations, such that:

$$P = \begin{pmatrix} 1 & \ldots & \rho_{1l} & \ldots & \rho_{1L} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \rho_{1L} & \ldots & \rho_{Ll} & \ldots & 1 \end{pmatrix}$$

We will denote the multivariate normal density of a variable $\mathbf{u} = (u_1, \ldots, u_L) \in \mathbb{R}^L$ with mean M and variance matrix Ω as

*email: mvargasbarrenechea@wiwi.uni-bielefeld.de; phone:+49 521 106 4917; fax:+49 521 106 2997; address: Institute of Mathematical Economics, Bielefeld University 33501, Bielefeld/Germany.

$\phi_L(\mathbf{u}; M, \Omega) = (2\pi)^{-L/2} |\Omega|^{-1/2} e^{-1/2\left[(\mathbf{u}-M)'\Omega^{-1}(\mathbf{u}-M)\right]}$, then

$$\begin{aligned}\Phi_L(\mathbf{w_i}; 0, R_i) = \int_{-\infty}^{w_{i,L}} \dots \int_{-\infty}^{w_{i,l}} \dots \int_{-\infty}^{w_{i,1}} \phi_L(\mathbf{u}; 0, R_i) du_1 \dots du_L = \\ = \int_{A_i} \phi_L(\mathbf{u}; 0, R_i) d\mathbf{u}\end{aligned} \tag{1}$$

where $A_i = [-\infty, w_{i,1}] \times \dots \times [-\infty, w_{i,l}] \times \dots \times [-\infty, w_{i,L}]$.

It's straightforward to prove that in the multivariate case the log-likelihood function is:

$$\ell(\beta, P|x) = \sum_{i=1}^{N} \log \Phi_L(\mathbf{w_i}; 0, R_i)$$

where $\beta = (\beta_1, \dots, \beta_l, \dots, \beta_L)'$, $\mathbf{w_i} = (w_{i,1}, \dots, w_{i,L})'$, $w_{i,l} = (2y_{i,l} - 1)x_{i,l}\beta_l$, $R_i = Q_i P Q_i$ and Q_i is a diagonal matrix $N \times N$ with diagonal $(2y_i - 1)$ and zeros in the other elements.

2 Derivatives

In this section we will find the analytical expressions for the first and second derivatives of the log-likelihood function. At the begining we will introduce some nomenclature, since $R_i = Q_i P Q_i$ we know that R_i it is a symmetric matrix with ones along the diagoinal:

$$R_i = \begin{pmatrix} 1 & \dots & r_{i,1l} & \dots & r_{i,1L} \\ \vdots & \ddots & \vdots & & \vdots \\ \vdots & & \vdots & \ddots & \vdots \\ r_{i,1L} & \dots & r_{i,Ll} & \dots & 1 \end{pmatrix}$$

now by reordering this matrix we obtain:

$$R_i^l = \left(\begin{array}{ccc|c} 1 & \dots & r_{i,1L} & r_{i,1l} \\ \vdots & \ddots & \vdots & \vdots \\ r_{i,1L} & \dots & 1 & r_{i,Ll} \\ \hline r_{i,1l} & \dots & r_{i,Ll} & 1 \end{array}\right) = \begin{pmatrix} R_{i,11}^l & R_{i,12}^l \\ R_{i,21}^l & 1 \end{pmatrix}$$

$$R_i^{kl} = \left(\begin{array}{ccc|cc} 1 & \dots & r_{i,1L} & r_{i,1k} & r_{i,1l} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ r_{i,1L} & \dots & 1 & r_{i,Lk} & r_{i,Ll} \\ \hline r_{i,1k} & \dots & r_{i,Lk} & 1 & r_{i,kl} \\ r_{i,1l} & \dots & r_{i,Ll} & r_{i,kl} & 1 \end{array}\right) = \begin{pmatrix} R_{i,11}^{kl} & R_{i,12}^{kl} \\ R_{i,21}^{kl} & R_{i,22}^{kl} \end{pmatrix}$$

Proposition 1.

$$\frac{\partial \ell(\beta, P|\mathrm{x})}{\partial \beta_l} = \sum_{i=1}^{N} \frac{\phi(w_{i,l}; 0, 1)\Phi_{L-1}(\mathbf{w_{i,-1}}; M^l, \Omega^l)(2y_{i,l} - 1)x_{i,l}}{\Phi_L(\mathbf{w_i}; 0, R_i)} \tag{2}$$

where $M_i^l = R_{i,12}^l w_{il}$, $\Omega_i^l = R_{i,11}^l - R_{i,12}^l R_{i,21}^l$ *and* $\mathbf{w_{i,-l}} = (w_{i,1}, \dots,$ $\dots, w_{i,l-1}, w_{i,l+1}, \dots, w_{i,L})$.

Proof. By known facts

$$\Phi_L(w_i, R_i) = \int_{A_i} \phi_L(\mathbf{u}; 0, R_i) du = \int_{A_i} \phi_{L-1}(\mathbf{u}_{-l}; M_l, \Omega_i^l)\phi(u_l; 0, 1) d\mathbf{u} =$$

$$= \int_{A_i} \phi(u_l; 0, 1)\phi_{L-1}(\mathbf{u}_{-l}; M_l, \Omega_i^l) d\mathbf{u} = \int_{-\infty}^{w_{i,l}} \phi(u_l; 0, 1) \times \Phi_{L-1}(\mathbf{w_{i,-l}}; M_l, \Omega_i^l) du_l$$

where $M_l = R_{i,12}^l u_l$, $\mathbf{u}_{-l} = (u_1, \dots, u_{l-1}, u_{l+1}, \dots, u_L)$ and $\mathbf{w_{i,-l}} = (w_{i,1}, \dots$ $, \dots, w_{i,l-1}, w_{i,l+1}, \dots, w_{i,L})$, then

$$\frac{\partial \Phi(w_i; 0, R_i)}{\partial w_{i,l}} = \phi_1(w_{i,l}; 0, 1)\Phi_{L-1}(\mathbf{w_{i,-l}}; M_i^l, \Omega^l) \tag{3}$$

because $w_{i,l} = (2y_{i,l} - 1)x_{i,l}\beta_l$ we have that

$$\frac{\partial \Phi(\mathbf{w_i}; 0, R_i)}{\partial \beta_l} = \phi_1(w_{i,l}; 0, 1)\Phi_{L-1}(\mathbf{w_{i,-l}}; M_i^l, \Omega^l)(2y_{i,l} - 1)x_{i,l} \tag{4}$$

where $M_i^l = R_{i,12}^l w_{il}$, $\Omega_i^l = R_{i,11}^l - R_{i,12}^l R_{i,21}^l$. By using the last result and the definition of $\ell(\beta, P|x)$ we find the wished result

$$\frac{\partial \ell(\beta, P|x)}{\partial \beta_l} = \sum_{i=1}^{N} \frac{\phi(w_{i,l}; 0, 1)\Phi_{L-1}(\mathbf{w_{i,-1}}; M^l, \Omega^l)(2y_{i,l} - 1)x_{i,l}}{\Phi_L(\mathbf{w_i}; 0, R_i)}$$

$\square$

Proposition 2.

$$\frac{\partial \ell(\beta, P|\mathrm{x})}{\partial \rho_{kl}} = \sum_{i=1}^{N} \frac{\phi_2(w_{i,k}, w_{i,l}; 0, R_{i,22}^{kl})\Phi_{L-2}(\mathbf{w_{i,-kl}}; M_i^{kl}, \Omega_i^{kl})}{\Phi(\mathbf{w_i}, R_i)} \times$$
$$\times (2y_{i,k} - 1)(2y_{i,l} - 1) \tag{5}$$

where $M_i^{kl} = R_{i,12}^{kl}(R_{i,22}^{kl})^{-1}(w_{i,l}, w_{i,k})'$, $\Omega_i^{kl} = R_{i,11}^{kl} - R_{i,12}^{kl}(R_{i,22}^{kl})^{-1}R_{i,21}^{kl}$ *and* $\mathbf{w_{i,-kl}} = (w_1, \dots, w_{k-1}, w_{k+1}, \dots, w_{l-1}, w_{l+1}, \dots w_L)$

Proof. By known facts

$$\Phi_L(w_i, R_i) = \int_{A_i} \phi_L(\mathbf{u}; 0, R_i) d\mathbf{u} =$$
$$= \int_{A_i} \phi_{L-2}(u_{-kl}; 0, R_{i,11}^{kl})\phi_2(u_l, u_k; M_{i,kl}, \Omega_{i,kl}) d\mathbf{u} =$$
$$= \int_{A_{i,-kl}} \left[\phi_{L-2}(u_{-kl}; 0, R_{i,11}^{kl}) \int_{-\infty}^{w_{i,k}} \int_{-\infty}^{w_{i,l}} \phi_2(u_l, u_k; M_{i,kl}, \Omega_{i,kl}) du_l du_k\right] d\mathbf{u_{-kl}} =$$
$$= \int_{A_{i,-kl}} \left[\phi_{L-2}(u_{-kl}; 0, R_{i,11}^{kl})\Phi_2(w_{il}, w_{ik}; M_{i,kl}, \Omega_{i,kl})\right] d\mathbf{u_{-kl}} \tag{6}$$

where

$$
\begin{aligned}
A_{i,-kl} &= [-\infty, w_{i,1}] \cdots [-\infty, w_{i,k-1}][-\infty, w_{i,k+1}] \cdots \\
&\cdots [-\infty, w_{i,l-1}][-\infty, w_{l+1,l}] \cdots [-\infty, w_{i,L}] \\
\mathbf{u_{-kl}} &= (u_1, \ldots, u_{k-1}, u_{k+1}, \ldots, u_{l-1}, u_{l+1}, \ldots u_L) \\
M_{i,kl} &= R^{kl}_{i,21}(R^{kl}_{i,11})^{-1}\mathbf{u_{-kl}} \\
\Omega_{i,kl} &= R^{kl}_{i,22} - R^{kl}_{i,21}(R^{kl}_{i,11})^{-1}R^{kl}_{i,12}
\end{aligned}
$$

Without loss of generality we will re-formulate $M_{i,kl}$ and $\Omega_{i,kl}$,

$$
M_{i,kl} = \begin{pmatrix} a_{1i} \\ a_{2i} \end{pmatrix}
$$

and

$$
\Omega_{i,kl} = \begin{pmatrix} 1 & r_{i,kl} \\ r_{i,kl} & 1 \end{pmatrix} - \begin{pmatrix} b_{11i} & b_{12i} \\ b_{12i} & b_{22i} \end{pmatrix}
$$

Since we know that only $\Phi_2(w_{il}, w_{ik}; M_{i,kl}, \Omega_{i,kl})$ depends on ρ_{kl}, we will just analyze the derivative of the second expression of the integrand in (6).

We know that

$$
\Phi_2(w_{il}, w_{ik}; M_{i,kl}, \Omega_{i,kl}) = \Phi_2^*(w^*_{ik}, w^*_{il}; \rho^*_{i,kl})
$$

where $w^*_{ik} = \dfrac{w^*_{ik} - a_{1i}}{\sqrt{1 - b_{11i}}}$, $w^*_{il} = \dfrac{w^*_{il} - a_{2i}}{\sqrt{1 - b_{22i}}}$, $\rho^*_{i,kl} = \dfrac{r_{i,kl} - b_{12i}}{\sqrt{(1 - b_{11i})(1 - b_{22i})}}$

$$
\Phi_2^*(w^*_{il}, w^*_{ik}; \rho^*_{i,kl}) = \int_{-\infty}^{w^*_{i,l}} \int_{-\infty}^{w^*_{i,k}} \phi_2^*(u_k, u_l; \rho^*_{i,kl}) du_k du_l
$$

with

$$
\phi_2^*(u_k, u_l; \rho^*_{i,kl}) = \frac{e^{-1/2\left[u_l^2 + u_k^2 - 2\rho^*_{i,kl} u_l u_k\right]/(1 - (\rho^*_{i,kl})^2)}}{2\pi\sqrt{1 - (\rho^*_{i,kl})^2}}
$$

Notice that the last expression is the density of the standard bivariate normal distribution, then the limits w^*_{il}, w^*_{ik} and the correlation coefficient $\rho^*_{i,kl}$ are obtained by normalization using the mean $M_{i,kl}$ and the variance matrix $\Omega_{i,kl}$, notice too that only $\rho^*_{i,kl}$ depends on ρ_{kl}.

Then

$$
\frac{\partial \Phi_2(w_{il}, w_{ik}; M_{kl}, \Omega_{kl})}{\partial \rho_{kl}} = \frac{\partial \Phi_2^*(w^*_{il}, w^*_{ik}; \rho^*_{kl})}{\partial \rho^*_{kl}} \times \frac{\partial \rho^*_{kl}}{\partial r_{i,kl}} \times \frac{\partial r^*_{i,kl}}{\partial \rho_{kl}} \tag{7}
$$

now by using the

$$
\frac{\partial \Phi_2^*(w^*_{il}, w^*_{ik}; \rho^*_{kl})}{\partial \rho^*_{kl}} = \phi_2^*(w^*_{il}, w^*_{ik}; \rho^*_{kl}) \tag{8}
$$

(see Greene [1], pp 850) it is straightforward prove that

$$\frac{\partial\Phi_2(w_{il},w_{ik};M_{i,kl},\Omega_{i,kl})}{\partial\rho_{kl}} = \phi_2(w_{il},w_{ik};M_{kl},\Omega_{kl})(2y_{i,k}-1)(2y_{i,l}-1) \quad (9)$$

By using the last result in (6), we obtain that

$$\begin{aligned}
&\frac{\partial\Phi_L(w_i,R_i)}{\partial\rho_{kl}} = \\
&= \int_{A_{i,-kl}} \left[\phi_{L-2}(u_{-kl};0,R^{kl}_{i,11})\phi_2(w_{il},w_{ik};M_{kl},\Omega_{kl})(2y_{i,k}-1)(2y_{i,l}-1)\right] d\mathbf{u_{-kl}} = \\
&= \int_{A_{i,-kl}} \phi_L(\mathbf{u};0,R_i)(2y_{i,k}-1)(2y_{i,l}-1)d\mathbf{u_{-kl}} = \\
&= \phi_2(w_{il},w_{ik};0,R^{kl}_{i,22})\Phi_{L-2}(\mathbf{w_{i,-kl}};M^{kl}_i,\Omega^{kl}_i)(2y_{i,k}-1)(2y_{i,l}-1) \quad (10)
\end{aligned}$$

where

$$M^{kl}_i = R^{kl}_{i,12}(R^{kl}_{i,22})^{-1}(w_{i,l},w_{i,k})'$$

and

$$\Omega^{kl}_i = R^{kl}_{i,11} - R^{kl}_{i,12}(R^{kl}_{i,22})^{-1}R^{kl}_{i,21}$$

.

Finally using the last result and the definition of $\ell(\beta,P|x)$, we have that

$$\frac{\partial\ell(\beta,P|x)}{\partial\rho_{kl}} = \sum_{i=1}^{N}\frac{\phi_2(w_{i,k},w_{i,l};0,R^{kl}_{i,22})\Phi_{L-2}(\mathbf{w_{i,-kl}};M^{kl}_i,\Omega^{kl}_i)}{\Phi(\mathbf{w_i},R_i)}\times$$
$$\times\,(2y_{i,k}-1)(2y_{i,l}-1)$$

□

3 Conclusions

The first derivatives of the log likelihood function for the multivariate probit are analytical expressions and without considering the integral of the function it is just necessary to calculate the integrals with one order less to obtain these derivatives.

References

[1] W.H. Greene. Econometric Analysis (4th ed.). Upper Saddle River, 2000.