Discussion Paper No. 588

CONVERGENCE OF INTEGRALS ENCOUNTERED IN DICHOTOMOUS DEPENDENT VARIABLE PROBLEMS*

by

Peter E. Rossi**

March 1984

---

*Financial support from the J. L. Kellogg foundation is gratefully acknowledged.

**J. L. Kellogg Graduate School of Management, Northwestern University, 2001 N. Sheridan Road, Evanston, Illinois 60201.

In the Bayesian analysis of dichotomous dependent variable models (c.f., Zellner and Rossi [1984]), the posterior expectation of various functions of the model parameters must be calculated. For example, the usual Bayes estimators of model parameters are the posterior means of those parameters. Integrals of the form $I = \int h(\underline{\beta})p(\underline{\beta}|D)d\underline{\beta}$, are considered where $p(\underline{\beta}|D)$ is the posterior density of $\underline{\beta}$. Using Bayes theorem, I can be rewritten as

$$I = [\int h(\underline{\beta})p(\underline{\beta})\ell(\underline{\beta}|D)d\underline{\beta}] / [\int p(\underline{\beta})\ell(\underline{\beta}|D)d\underline{\beta})]$$

where $\ell(\underline{\beta}|D)$ is the likelihood function and $p(\underline{\beta})$ is the prior density. In some cases, a diffuse prior $(p(\underline{\beta}) \propto k)$ is employed and the convergence of I depends on the convergence of $\int \ell(\underline{\beta}|D)d\underline{\beta}$ and the behavior of $h(\underline{\beta})$ function. In this note, necessary and sufficient conditions for the convergence of $\int \ell(\underline{\beta}|D)d\underline{\beta}$ are derived for the logit model.

The logit model is a sequence of binomial trails.

$$Y_i = \begin{cases} 1 & \text{with prob. } P_i \\ 0 & \text{with prob. } (1 - P_i) \end{cases} \qquad i = 1, \ldots, n$$

$$\text{with } P_i = 1/(1 + \exp(-\underline{x}_i'\underline{\beta}))$$

The likelihood function for this model can be written

$$\ell(\underline{\beta}|D) = \prod_{i=1}^{n} P_i^{Y_i}(\underline{\beta})(1 - P_i(\underline{\beta}))^{1-Y_i}$$

We seek conditions under which $I' = \int_{\mathbb{R}^k} \ell(\underline{\beta}|D)d\beta$ converges. The convergence of I' depends on the behavior of $\ell(\underline{\beta}|D)$ for large (in the sense of $\|\underline{\beta}\|$) $\underline{\beta}$ vectors. Not only must $\ell(\underline{\beta}|D)$ damp off to zero but it also must do so at a fast enough rate. In the logit model, the tails of $\ell(\underline{\beta}|D)$ will either approach zero or a constant limit at an exponential rate. Thus, we must determine under what conditions $\lim_{\|\underline{\beta}\|\to\infty} \ell(\underline{\beta}|D) = k$ where k is a non-zero constant. If there are directions along which $\lim_{\|\underline{\beta}\|\to\infty} \ell(\underline{\beta}|D) \neq 0$, then the integral of $\ell(\underline{\beta}|D)$ over $\mathbb{R}^k$ will not be defined. In the section below, precise mathematical arguments are developed to support these intuitive arguments. In addition, our convergence conditions are compared to the conditions for the existence of a unique maximum likelihood estimator developed by Silvapulle (1981) and to the Zellner-Rossi sufficient conditions for the two-dimensional case.

<u>Limiting Behavior of $\ell(\beta|D)$</u>

For the logit model, we choose $P_i = 1/(1 + \exp(-\underline{x}_i'\underline{\beta}))$. The tail behavior of the likelihood function depends on the limiting properties of each of the P_i.

Consider some increasing sequence of $(\underline{\beta})$ vectors, $\{\underline{\beta}^i\}$ with $\lim_{i\to\infty} \|\underline{\beta}^i\| = \infty$. All that is necessary for $\ell(\cdot)$ to damp off to zero along this sequence is for one of the P_i to tend to 0 or 1 depending on the value of Y_i. Ordering the observations so that $Y_i = 0$ for i= 1,...,m and $Y_i = 1$ for i= m+1,...,n, we write the likelihood function as

$$\ell(\underline{\beta}) = P_1 \cdots P_m \; (1 - P_{m+1}) \cdots (1 - P_n) \; .$$

Since each term (P_i or $(1-P_i)$) in the product which makes up $\ell(\underline{\beta})$ is bounded, only one term is necessary to drive $\ell(\cdot)$ to zero. Thus, in order to produce a non-zero limit of $\ell(\cdot)$ along a given sequence of $\underline{\beta}$'s, we must drive each of the terms in $\ell(\underline{\beta})$ to some non-zero number. The following lemma determines the possible limiting values of $\ell(\cdot)$ and establishes that it is only necessary to inspect sequences of $\underline{\beta}$'s along rays extended from the origin.

Lemma 1. Consider any increasing sequence of $\underline{\beta}$'s, $\{\underline{\beta}^i\}$ $\lim_{i\to\infty} \|\beta^i\| = \infty$.

(a) $\lim_{i\to\infty} \ell(\underline{\beta}^i) = 0$ or 1 if the design matrix is of full rank.

(b) $\{\underline{\beta}^{1i}\}$ is a sequence such that $\lim_{i\to\infty} \ell(\underline{\beta}^{1i}) \neq 0$ iff $\exists$ a sequence $\{\underline{\beta}^{2i}\}$ along a ray $(\underline{\beta}^{2(i+1)} = k\underline{\beta}^{2i})$ with $\lim_{i\to\infty} \ell(\underline{\beta}^{2i}) \neq 0$.

Proof of Lemma 1:

(a) Since $\lim_{i\to\infty} \|\underline{\beta}^i\| = \infty$, $\lim_{i\to\infty} \underline{x}_i'\underline{\beta} = \pm\infty$ or 0.

$$\lim_{i\to\infty} P_j(\underline{\beta}^i) = \lim_{i\to\infty} \frac{1}{(1 + e^{-\underline{x}_i'\underline{\beta}^i})} = \begin{cases} 1 & \text{if } \lim_{t\to\infty} \underline{x}_i'\underline{\beta} = +\infty \\ 0 & \text{if } \lim_{i\to\infty} \underline{x}_i'\underline{\beta} = -\infty \\ .5 & \text{if } \lim_{i\to\infty} \underline{x}_i\underline{\beta} = 0 \end{cases}$$

Each P_j tends to 1, 0, or .5.

If all P_j tend to .5, then $\lim_{i\to\infty} \ell(\underline{\beta}^i) = (.5)^n$ However, in this case, the sequence $\{\underline{\beta}^i\}$ must approach a nontrivial solution to $X\,\underline{\beta} = \underline{0}$ which does not exist if X is of full rank.

If all P_j tend to 1 $(j= 1,\ldots,m)$ and 0 $(j= m+1,\ldots,n)$, then $\lim_{i\to\infty} \ell(\underline{\beta}^i) = 1$.

If any P_j tends to 0 $(j= 1,\ldots,m)$ or 1 $(j= m+1,\ldots,n)$, then $\lim_{i\to\infty} \ell(\underline{\beta}^i) = 0$ since $\ell(\cdot)$ is bounded.

(b) Consider a sequence $\{\underline{\beta}^{1i}\}$ with $\lim_{i\to\infty} \ell(\underline{\beta}^{i}) = 1$. For large enough i, there is a $\underline{\beta}^{1i} = \underline{\beta}^{*}$ which satisfies

(1a) $\qquad \underline{x}_j'\underline{\beta}^{*} > 0 \qquad j=1,\ldots,m$

(1b) $\qquad \underline{x}_j'\underline{\beta}^{*} < 0 \qquad j = m+1,\ldots,n.$

Construct the sequence $\{\underline{\beta}^{2i} = k\underline{\beta}^{*} \quad k=1,2,\ldots\}$. Along this sequence, $\lim_{i\to\infty} \ell(\underline{\beta}^{2i}) = 1$.

Integral Convergence

In order to better understand integral convergence conditions, we examine first the situations in which $\int \ell(\underline{\beta})d\underline{\beta}$ will diverge. $\int \ell(\underline{\beta})d\underline{\beta}$ will diverge when there are tail regions where $\ell(\underline{\beta})$ does not damp off to zero. This occurs when $\ell(\underline{\beta})$ tends to 1 along some ray.

Lemma 2. If there is a direction, $\underline{\beta}^{*}$, with $\lim_{i\to\infty} \ell(k\underline{\beta}^{*}) = 1$, then $\int \ell(\underline{\beta})d\underline{\beta})$ will diverge. We note that $\ell(\cdot)$ approaches either 0 or 1 along any increasing sequence (see Lemma 1).

Proof of Lemma 2. Take a direction, $\underline{\beta}^{*}$ for which $\lim_{k\to\infty} \ell(k\underline{\beta}^{*}) = 1$. By definition, we can find a K such that for all $k \geq K$.

$$\left|\ell(k\underline{\beta}^{*}) - 1\right| < \varepsilon/2$$

Now let us use the continuity of $\ell(\cdot)$ to choose a sequence of non-overlapping δ neighborhoods of points along the ray in the direction of $\underline{\beta}^{*}$.

By continuity of $\ell(\cdot)$, we can pick a sequence of $\{k^i, \ i = 1,\ldots,n\}$ such that $|\ell(\underline{\beta}) - \ell(k^i\underline{\beta}^*)| < \varepsilon/2$ for all $\underline{\beta} \in A_i$ where $A_i = \{\underline{\beta}: \|\underline{\beta} - k^i\underline{\beta}^*\| < \delta_i\}$ and $A_i \cap A_j = \emptyset$ for $i \neq j$

$$\int \ell(\underline{\beta})d\underline{\beta} \geq \sum_i \int_{A_i} \ell(\underline{\beta})d\underline{\beta} \geq \sum_i (1 - \varepsilon)\delta_i$$

By increasing the number of terms in the sequence of k^i, we can make $\sum_i (1 - \varepsilon)\delta_i$ as large as possible.

If the tails of $\ell(\cdot)$ damp off to zero in all directions, is this sufficient to ensure convergence of $\ell(\cdot)$?

<u>Theorem 1.</u> $\int \ell(\underline{\beta})d\underline{\beta}$ is finite under the following conditions:

(i) The design matrix, $\underset{n\times k}{X}$, is of full column rank.

(ii) Let $A = \{\underline{\beta}: \underline{x}_i'\underline{\beta} > 0 \quad i = 1,\ldots,m\}$ and

$B = \{\underline{\beta}: x_i'\underline{\beta} < 0 \quad i = m+1,\ldots,n\}$ then $A \cap B = \emptyset$.

<u>Proof of Theorem 1:</u> If $\exists$ $\underline{\beta}^*$ such that $X\underline{\beta}^* = 0$ and $\underline{\beta}^* \neq \underline{0}$ then $\lim_{k\to\infty} (k\underline{\beta}^*) = (.5)^n$. Lemma 2 tells us that under those conditions $\int \ell(\underline{\beta})d\underline{\beta}$ will diverge. If X is of full rank, we cannot find this $\underline{\beta}$. If $\exists$ $\underline{\beta}^* \in A \cap B$, then

$$\lim_{k\to\infty} P_i(k\underline{\beta}^*) = 1/(1 + \exp(-k\underline{x}'\underline{\beta}^*)) = 1 \quad i = 1,\ldots,m$$

and

$$\lim_{k\to\infty} P_i(k\underline{\beta}^*) = 0 \quad i = m+1,\ldots,n.$$

Again, by Lemma 2 $\int \ell(\underline{\beta})d\underline{\beta}$ diverges. Lemma 1 indicates that it is sufficient to examine only those directions along rays.

If $\lim_{k\to\infty} \ell(k\underline{\beta}^*) = 0$ for all $\underline{\beta}^*$, will $\int \ell(\underline{\beta})d\beta$ converge? To establish that this is true, we convert to polar coordinates--that is, we first pick a direction, integrate along that ray, and then integrate over all possible directions

$$\int \ell(\underline{\beta})d\underline{\beta} = \int_S \int_0^\infty \ell(k\underline{\beta}^*)dkd\underline{\beta}^*$$

where

$$S = \{\underline{\beta}^* : \beta_1^{*2} + \dots + \beta_k^{*2} = 1\}$$

Fix $\underline{\beta}^*$ and find a term in $\ell(\cdot)$ which approaches zero. This will occur either if $P_i \to 0$ for any i between 1 and m or if $P_i \to 1$ for any i between m+1 and n. $P_i \to 1$ only if $\underline{x}_i'\underline{\beta}^* > 0$, and $P_i \to 1$ only if $\underline{x}_i'\underline{\beta} > 0$. Condition (i) ensures that we can find a $\underline{x}_i'\underline{\beta}^*$ which is not zero. Condition (ii) implies that for every direction $\underline{\beta}^*$ some term $(P_i$ or $(1 - P_i))$ will approach zero.

Define $r(\underline{\beta}^*) = \max[-\underline{x}_i'\underline{\beta}^*, \dots, -\underline{x}_m'\underline{\beta}^*, \underline{x}_{m+1}'\underline{\beta}^*, \dots, \underline{x}_n'\underline{\beta}^*]$. $e^{-r(\underline{\beta}^*)}$ is the rate at which $\ell(k\underline{\beta}^*)$ approaches zero as k increases.

Define $\bar{r} = \inf_{\underline{\beta}^* \varepsilon S} \{r(\underline{\beta}^*)\}$. Conditions (i) and (ii) assure that $r(\underline{\beta}^*) > 0$. Since $r(\underline{\beta}^*)$ is a continuous function defined on the compact set, $r(\underline{\beta}^*)$ attains its infinum, $\bar{r}$. Therefore, $\bar{r} > 0$.

$$\ell(k\underline{\beta}^*) = P_1(k\underline{\beta}^*) \dots P_m(k\underline{\beta}^*)\left(1 - P_{m+1}(k\underline{\beta}^*)\right) \dots \left(1 - P_n(k\underline{\beta}^*)\right)$$

$$= \prod_{i=1}^{m} \frac{1}{[1 + \exp(-kx_i'\underline{\beta}^*)]} \cdot \prod_{i=m+1}^{n} \frac{1}{[1 + \exp(kx_i'\underline{\beta}^*)]}$$

$$\leq \frac{1}{1 + \exp(k\bar{r})}$$

Thus,

$$\int_S \int_0^\infty \ell(k\underline{\beta}^*) \leq \int_S \int_0^\infty \frac{1}{1 + \exp(k\bar{r})} dkd\underline{\beta}^*$$

$$\leq \int_S \int_0^\infty e^{-k\bar{r}} dkd\underline{\beta}^* = \frac{1}{\bar{r}} \int_S d\underline{\beta}^* < \infty \quad \text{as S is bounded.}$$

The Two Dimensional Case

In this section, we examine the conditions (i) and (ii) required for integral convergence in the two-dimensional case. We also relate these conditions to the sufficient conditions developed by Zellner-Rossi (1984).

For $k = 2$, condition (ii) can be rewritten as the system of inequalities

$$\begin{aligned}
\beta_0 + \beta_1 x_1 &> 0 \\
&\vdots \\
\beta_0 + \beta_1 x_m &> 0 \\
\beta_0 + \beta_1 x_{m+1} &< 0 \\
&\vdots \\
\beta_0 + \beta_1 x_n &< 0
\end{aligned}$$

The likelihood function is integrable only if there is no non-zero $\underline{\beta}$ which satsifies the above inequalities. In the two dimensional case, two restrictions on the x_i ensure that there is no solution to the system of inequalities.

(i) $\max(x_1,\ldots,x_m) > \min(x_{m+1},\ldots,x_n)$

and

(ii) $\min(x_1,\ldots,x_m) < \max(x_{m+1},\ldots,x_n)$

These conditions require that there is some "overlap" between the x values for which Y = 1 and for which Y = 0. If there is no overlap, then the likelihood function will increase and asymptote to one along some direction from the origin. In Zellner and Rossi (1984), a simple, three observation sample is considered (m=1, n=3).

i	1	2	3
Y_i	0	1	1
X_i	-1	0	1

with $P_i = 1/\left(1 + \exp(-\beta_0 - \beta_1 x_i\right)$

By increasing β_1, the logit curve bends to fit the sample proportions. The key to understanding the example is the monotonic increase in x_i. Any ordered dataset with monotonically increasing x_i will generate a non-integrable likelihood function.

Relationship to Zellner-Rossi Sufficient Conditions

Zellner and Rossi (1984) develop sufficient conditions for integral convergence in the case $k = 2$.

Zellner-Rossi conditions:

(ZR1) $$\sum_{i=1}^{m} x_i > \sum_{i=1}^{m+1} x_{(i)}$$

(ZR2) $$\sum_{i=1}^{m} x_i < \sum_{i=1}^{m-2} x_{(n-i)}$$

Here $x_{(i)}$ is the ith order statistic of the sample $(x_1,\ldots,x_n)$.

Theorem 2. (ZR1) and (ZR2) imply conditions (i) and (ii) above.

Proof of Theorem 2. Let us look at the contrapositive. Suppose $\max(x_1,\ldots,x_m) < \min(x_{m+1},\ldots,x_n)$. Then

$$\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} x_{(i)} \Rightarrow \sum_{i=1}^{m} x_{(i)} + x_{(m+1)} > \sum_{i=1}^{m} x_i$$

Suppose $\min(x_1,\ldots,x_m) > \max(x_{m+1},\ldots,x_n)$. Then

$$\sum_{i=0}^{m-2} x_{(n-i)} = \sum_{i=1}^{m} x_i - \min(x_1,\ldots,x_m) \Rightarrow \sum_{i=0}^{m-2} x_{(n-i)} < \sum_{i=1}^{m} x_i$$

The Zellner and Rossi conditions are not equivalent to our conditions (i) and (ii) above. The Zellner-Rossi conditions are sufficient but not necessary. Consider the following example of a dataset which satisfies (i) and (ii) but not (ZR1) and (ZR2).

i	1	2	3	4
Y_i	0	0	1	1
X_i	1	2	1.5	3

m=2, n=4

$$\sum_{i=1}^{2} x_i = 3 < 1 + 2 + 1.5 = 4.5 \quad \text{(violates (ZR1))}$$

and

$$\max(1,\ 2) > \min(1.5,\ 3) \quad \text{(condition (i))}$$

$$\min(1,\ 2) < \max(1.5,\ 3) \quad \text{(condition (ii))}$$

Relationship to Conditions Guaranteeing the Existence of MLE

Silvapulle (1981) investigates conditions under which the maximum likelihood estimator of slope coefficient vector, $\underline{\beta}$, exists and is unique. Silvapulle considers the general model where $\text{Prob}(Y_i = 1) = G(\underline{x}_i\underline{\beta})$. For the logit case, the maximum likelihood estimator of $\underline{\beta}$ exists and is unique if and only if the following condition is met.

Let S, F be the relative interiors* of the convex cones generated by $\underline{x}_1,\ldots,\underline{x}_m$ (the values of independent variables for the first m in ordered observations with $Y_i = 0$) and $\underline{x}_{m+1},\ldots,\underline{x}_n$, respectively.

$$S = \{\underline{\beta} : \underline{\beta} = \sum_{i=m+1}^{n} k_i \underline{x}_i,\quad k_i > 0\}$$

*In the sense of Rockafellar (1972).

$$F = \{\underline{\beta}: \underline{\beta} = \sum_{i=1}^{m} k_i \underline{x}_i, \quad k_i > 0\}$$

The MLE of $\underline{\beta}$ exists and is unique if and only if

(S1) $\quad S \cap F \neq \emptyset$

(S1) ensures that there is some overlap between the cone generated by observations with $Y_i = 0$ and that cone generated by observations with $Y_i = 1$. This is exactly the same intuition behind our conditions in Theorem 1 which ensure the existence of moments of the posterior distribution. We will now show the equivalence of our integral convergence conditions with Silvapulle's conditions.

Theorem 3. $S \cap F \neq \emptyset$ if and only if $A \cap B = \emptyset$ where S, F are defined above and A, B are defined in Theorem 1.

Proof of Theorem 3. Recall that

$$A = \{\underline{\beta}^*: \underline{x}_i'\underline{\beta}^* > 0 \quad i=1,\ldots,m\}$$

and

$$B = \{\underline{\beta}^*: \underline{x}_i'\underline{\beta}^* < 0 \quad i=m+1,\ldots,n\}$$

A is the polar cone* of F. B is the negative of the polar cone of S. We can rewrite the conditions as follows

---

*The polar cone of a set C, denoted $C^o \equiv \{x^*: \underline{x}^{*\prime}\underline{x} \leq 0, x \in C\}$.

$$S \cap F \neq \emptyset \quad \text{iff} \quad (-S^{o}) \cap F^{o} = \emptyset$$

In Figure 1a, we illustrate the case where $S \cap F = \emptyset$ and in Figure 1b the case of non-null intersection is shown.

$A \cap B = \emptyset$ implies there is no solution to the system

$$(*) \qquad \underset{n\times k}{\bar{X}}\,\underline{\beta} \geq \underset{n\times 1}{\underline{0}} \quad \text{with} \quad \bar{X}' = \{ X'_{(m)} \,\vdots\, -X'_{(n-m)} \}$$
$$X'_{(m)} = [\underline{x}_1, \ldots, \underline{x}_m],\ X'_{(n-m)} = [\underline{x}_{m+1}, \ldots, \underline{x}_n].$$

$S \cap F \neq \emptyset$ implies there exists $\underset{n\times 1}{\underline{k}} \geq \underline{0}$ such that

$$(**) \qquad \bar{X}'\underline{k} = \underline{0}$$

To show equivalence, we must demonstrate that either there exists a solution to the system

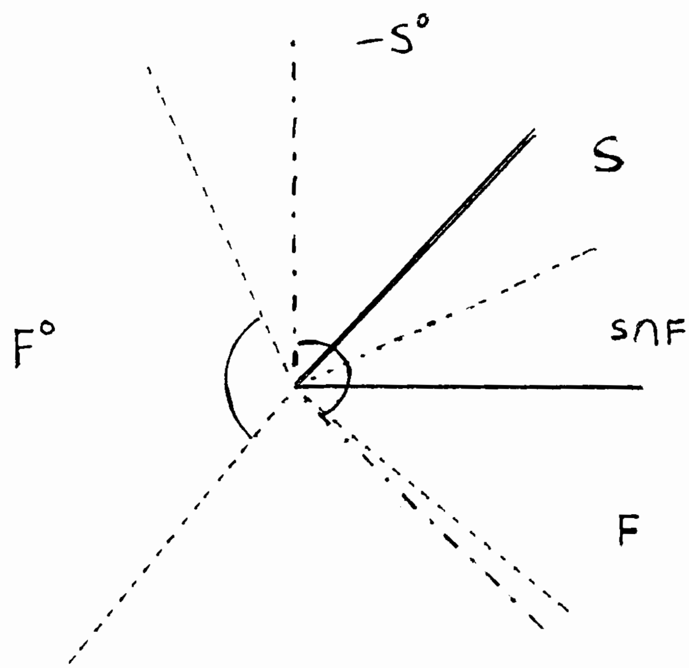
$$(*') \qquad \bar{X}\,\underline{\beta} < \underline{0}$$

or that (**) obtains, but never both. Clearly, (*') and (**) cannot both be true. If $\underline{\tilde{\beta}}$ satisfies (*') and $\underline{\tilde{k}}$ satisfies (**), $\underline{\tilde{k}}'\bar{X}\,\underline{\tilde{\beta}} < 0$ and $\underline{\tilde{k}}'X\,\underline{\tilde{\beta}} = 0$. We must show that if there is no $\underline{\beta}$ which solves (*') then there exists $\underline{k} \geq \underline{0}$ such that (**) obtains. We use a version of Farkas' Lemma.

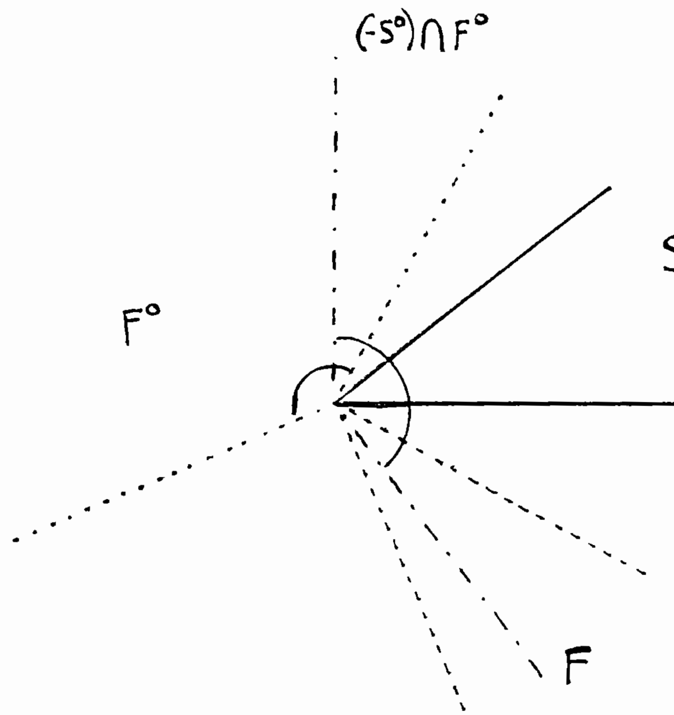
Figure 1.

Illustration of Equivalence of Integral Convergence Conditions

a. $S \cap F \neq \emptyset$



b. $S \cap F = \emptyset$



Farkas' Lemma. $\underline{b}'\underline{x} \geq 0$ is a consequence of $A\,\underline{x} \leq \underline{0}$ if and only if there exists $\underline{k} \geq \underline{0}$ such that $A'\underline{k} = \underline{b}$.

Proof of Lemma:. See Rockafellar, Corollary 22.3.1, pp. 200.

If there is no $\underline{\beta}$ satisfying $X\,\underline{\beta} < \underline{0}$, then there is no negative number, δ, such that

$$\bar{X}\,\underline{\beta} \leq \delta\,\underline{\iota} \quad \text{for all } \underline{\beta}\ , \text{ here } \iota' = (1, \ldots, 1)\ .$$

Let $\underline{x}' = (\delta, \underline{\beta}')$, $\underline{b}' = (1, \underline{0}')$ and $A = [\underline{\iota}, -\bar{X}]$, then it can be concluded from Farkas' Lemma that

$$\underline{b}'\underline{x} = \delta \geq 0 \quad \text{holds for} \quad \delta\,\underline{\iota} + \bar{X}\,\underline{\beta} \geq \underline{0}$$

if and only if there exists $\underline{k} \geq \underline{0}$ such that $\underline{\iota}'\underline{k} = 1$ and $\bar{X}\,\underline{k} = \underline{0}$. Hence, we have found a $\underline{k}$ which solves (**).

Verification of Integral Convergence and Existence of MLE

The conditions which ensure integrability of the likelihood function can be verified by application of simple linear programming techniques. According to Theorem 3, we may use either Silvapulle's or our own conditions to verify integrability. It is most convenient to use the conditions in Theorem 2 since verification only involves determining if there exists a solution to a system of linear inequalities. Let us rewrite the conditions given in Theorem 2 in matrix form.

The likelihood function is not integrable if there exists a non-zero solution to the following set of inequalities

$$\bar{X}\underline{\beta} \leq \underline{0}$$

In order to apply the Simplex algorithm, we must reduce the system to a standard linear programming problem. The $\underline{\beta}$ vector is written as $\underline{\beta} = \underline{v} - \underline{u}$ where $\underline{v}, \underline{u} \geq 0$. The system of inequalities can be rewritten as

$$\bar{X}\underline{v} - \bar{X}\underline{u} \leq \underline{0}$$

To determine if there are non-zero solutions to the above system of inequalities, we introduce a vector of artificial variables to this set of inequalities and minimize the sum of artificial variables. If any of the artificial variables remain in the basis on termination, there are no $\underline{\beta}$ vectors satisfying $\bar{X}\underline{\beta} \leq \underline{0}$ except $\underline{0}$ and the integral converges.

We solve the linear program

$$\begin{aligned} &\text{minimize } \underline{\iota}'\underline{y} \\ &\text{s.t. } \bar{X}\underline{v} - \bar{X}\underline{u} + \underline{y} = \underline{0} \\ &\qquad \underline{v}, \underline{u}, \underline{y} \geq \underline{0} \end{aligned}$$

We initiate the linear program with nonzero $\underline{y}$ (artificial) variables. This procedure is generally termed (see, for example, Luenberger (1973)) phase I of the simplex algorithm.

References

Luenberger, David G. 1973. Introduction to Linear and Nonlinear Programming. Massachusetts: Addison-Wesley.

Rockafellar, R. T. 1970. Convex Analysis. Princeton: Princeton University Press.

Silvapulle, Mervyn J. 1981. "On the Existence of Maximum Likelihood Estimaters for the Binomial Response Models." Journal of The Royal Statistical Society, Series B, No. 43, pp. 310-313.

Zellner, Arnold and Peter E. Rossi. 1984. "Bayesian Analysis of Quantal Response Models." Journal of Econometrics, forthcoming.