

ロジットモデル構築における Weight of Evidence を用いた変数変換と欠損値処理方法の提案

(株) 金融工学研究所 木村和央*

1. はじめに

ロジットモデル構築にあたっては、限られた時間コストの中で、人的スキルによらずに、一定精度を確保したモデルを得たいという実務上の要請がある。大勢待 [1] は Weight of Evidence (WoE) を用いて効率的にデフォルト確率モデルが構築可能なことを明らかにした。本稿では、ソブリン (国) のデフォルト確率および格付該当確率モデルの構築における WoE (対数オッズ, Z スコア) のさらなる活用法を提案した。

2. データと分析の枠組み

対象国は世界銀行 (世銀) 加盟の 188 か国, 対象期間は 2001 から 2014 年 (14 年間) を採用した。データは, 世銀の WDI と WGI をベースとした¹。デフォルトは, 主要債権国会議 (パリクラブ) へ持ち込まれた年とした。格付は, 格付投資情報センター (R&I) が前年 12 月末に付与していた格付とした。データの収録時期を考慮し, $t-2$ 年の WDI と $t-3$ 年の WGI により, t 年のデフォルトと格付を推計した。

3. 外れ値処理と Yeo-Johnson 変換を用いた変数変換

WoE の定義は, サンプル全体と部分集合の (デフォルト率に関する) 対数オッズ比であり, 連続型変数の場合は離散化が必要であるが, 連続型変数のままロジットモデルを推計し, 得られた対数オッズ (Z スコア) を利用して WoE に変換してもよい。ただし, 連続型変数ゆえ, 外れ値の処理と変数変換の選択を考慮しなければならない。前者にはグリッドサーチ, 後者には Yeo-Johnson 変換 [2] を用いて対応した。

4. 重回帰モデルによる欠損値補完処理

ソブリンにおいては, 国により統計データの集計・把握に時間がかかること, あるいはそもそも集計対象でないといったことから, 欠損値の存在を前提にモデル構築を考えなければならない。ここでは, 前章の Z スコアが欠損値の部分につき, 重回帰モデルを繰り返し推計する手法 (単一代入法の 1 つ) を用いて補完した。重回帰モデルは 4 種類構築し, 欠損値パターン別に適用した。また, Z スコアの利点を生かし, 切片を除く係数パラメータは正值に制約させることで安定的な結果が得られた。

5. デフォルトと格付該当確率の同時推計

Z スコアの完全データをもとに, 最終的なロジットモデルの推計を実施した。デフォルト確率モデルと格付該当確率モデルは, 両モデルの対数尤度 (l_P, l_R) の加重平均値 ($l_T = a \cdot l_P + (1-a) \cdot l_R$; a は固定値) を最大化する制約条件付き最適化問題を解くことでパラメータの推計がなされた。

6. 最終結果と精度確認

一例として, 右表を参照。

参考文献

- [1] 大勢待 利明 (2008). デフォルト確率推定モデル作成における WOE 変換の役割とその利用方法. 2008SAS ユーザー総会論文集, 298-305.
- [2] Yeo, In-Kwon and Johnson, Richard (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, 87, 954-959.

| 最終モデル利用変数名 (p 値 < 0.005) | 説明力 (%) | |
|-----------------------------|---------|-----|
| | 共分散 | 自分散 |
| GDP 成長率 | -0.3 | 0.9 |
| 債務利息支払/歳入 | 4.7 | 2.7 |
| 消費者物価上昇率の 3 年 S.D. | 6.6 | 1.2 |
| 為替レート変化率の 3 年 S.D. | 2.9 | 0.9 |
| 1 人当たり GDP/世界平均 | 27.0 | 9.4 |
| GDP/世界合計 | 16.5 | 6.0 |
| 対外債務/GDP | 0.0 | 0.6 |
| 政府の有効性 (WGI) | 22.2 | 6.3 |
| 法の支配 (WGI) | 20.3 | 5.9 |

AR 値=0.773(±0.040), KS 値=0.640.

実績・推計格付間の順位相関係数=0.925.

*本稿の内容は筆者に属し, 所属組織の見解ではない。

¹WDI : World Development Indicators, WGI : The Worldwide Governance Indicators.