A Credit Risk Application of Multivariate Ordinal Regression Models using the R package mvord

Rainer Hirk, Kurt Hornik and Laura Vana
R/Finance 2018, Chicago
June 1st, 2018
Introduction and motivation

Overview R package mvord

Examples
Credit risk in a nutshell

- **Credit risk** is the risk of a loss arising from a failure (or default) of a counterparty to meet its contractual obligations (e.g., McNeil et al., 2015).

- **Credit ratings** are forward-looking opinions about the creditworthiness of an obligor.

- Credit ratings are **ordinal** assessments of credit risk and are primarily relevant for:
  - investors
  - regulators and legislators
  - issuers
  - financial institutions

- Credit ratings as well as (internally) estimated probabilities of default (PDs) are central measures of credit risk.
Motivation

- Two modeling approaches:
  - Credit rating models (e.g., Blume et al., 1998; Alp, 2013; Baghai et al., 2014)
  - Failure prediction models (e.g., Shumway, 2001; Tian et al., 2015)

- Correlated ordinal data
  - Multiple correlated ratings assigned by different raters to one firm at the same point in time.
  - For each rater, there is serial dependence over the years.

- There is need for a flexible model class that can handle correlated ordinal and binary data:
  1. Heterogeneity in the rating methodology
  2. Heterogeneity in the covariates
  3. Unbalanced panel of firms
Multivariate ordinal regression

- $i$ denotes the subject index (firm).
- $j$ denotes the multiple measurement index (rater).
- $Y_i = [Y_{ij}]_{j=1,\ldots,q}$ is a $(q \times 1)$ vector of correlated ordinal response variables which is observed together with covariates.
- The association between the $Y_i$’s is captured by a multivariate structure imposed on the latent variables $\tilde{Y}_i$:
  \[
  \tilde{Y}_{ij} = \beta_{0j} + x_{ij}^T \beta_j + \epsilon_{ij}, \quad \epsilon_i = [\epsilon_{ij}]_{j=1,\ldots,q} \sim F_{i,q}(0, \Sigma_i),
  \]
  where $F_{i,q}$ denotes the $q$-dimensional joint distribution of the errors $\epsilon_i$.
- For each $j$,
  \[
  Y_{ij} = r \iff \theta_{j,r-1} < \tilde{Y}_{ij} \leq \theta_{j,r}, \quad r \in \{1, \ldots, K_j\},
  \]
  where $-\infty = \theta_{j,0} < \theta_{j,1} < \cdots < \theta_{j,K_j-1} < \theta_{j,K_j} = \infty$ are response specific thresholds.
The R package **mvord** (Hirk et al., 2018) implements pairwise likelihood estimation in the class of multivariate ordinal regression models in a flexible framework.

- Several identifiability constraints are supported.
- Multivariate probit and logit links are implemented.
- The correlation between the variables is accounted for by different (covariate dependent) error structures:
  - `cor_general()`: $\text{corr}(\epsilon_{ik}, \epsilon_{il}) = \rho_{ikl}$,
  - `cor_equi()`: $\text{corr}(\epsilon_{ik}, \epsilon_{il}) = \rho_i$,
  - `cor_ar1()`: $\text{corr}(\epsilon_{ik}, \epsilon_{il}) = \rho_i^{\mid l-k \mid}$.
- Constraints on the threshold and coefficient parameters can be set.
- Category-specific coefficients are supported.
Data

- **Long-term issuer credit ratings** assigned by S&P, Moody’s and Fitch for US companies excluding the financial and utilities sectors;
  - S&P: AAA, AA, A, BBB, BB, B, CCC, CC
  - Fitch: AAA, AA, A, BBB, BB, B, CCC, CC, C
  - Moody’s: Aaa, Aa, A, Baa, Ba, B, Caa, Ca
- **Sources:** Compustat North America© Ratings File, Moody’s Default & Recovery Database©, Fitch Rating Services.

- Failure indicator: binary indicator set to one on occurrence of bankruptcy filing under Chapter 7 or Chapter 11, or default rating by CRAs in the one-year window following the rating observation;
- **Sources:** UCLA-LoPucki Bankruptcy Research Database, Mergent FISD ©.

- Covariates: financial ratios and market variables;
- **Pre-processing:** e.g., outlier removal by winsorization, removal of missing values.
- **Sources:** Compustat North America© Fundamentals Annual File, The Center for Research in Security Prices (CRSP).

- Period: 1999–2013
Long-term issuer credit ratings assigned by S&P, Moody’s and Fitch for US companies excluding the financial and utilities sectors;
  - S&P: Aaa, Aa, A, Baa, Ba, B, Caa, Ca
  - Fitch: AAA, AA, A, BBB, BB, B, CCC, CC, C
  - Moody's: Aaa, Aa, A, Baa, Ba, B, Caa, Ca

Sources: Compustat North America© Ratings File, Moody’s Default & Recovery Database©, Fitch Rating Services.

Failure indicator: binary indicator set to one on occurrence of bankruptcy filing under Chapter 7 or Chapter 11, or default rating by CRAs in the one year-window following the rating observation;
  - Sources: UCLA-LoPucki Bankruptcy Research Database, Mergent FISD©.
Long-term issuer credit ratings assigned by S&P, Moody’s and Fitch for US companies excluding the financial and utilities sectors;
- S&P: AAA, AA, A, BBB, BB, B, CCC, CC
- Fitch: AAA, AA, A, BBB, BB, B, CCC, CC, C
- Moody’s: Aaa, Aa, A, Baa, Ba, B, Caa, Ca
- Sources: Compustat North America© Ratings File, Moody’s Default & Recovery Database©, Fitch Rating Services.

Failure indicator: binary indicator set to one on occurrence of bankruptcy filing under Chapter 7 or Chapter 11, or default rating by CRAs in the one year-window following the rating observation;
- Sources: UCLA-LoPucki Bankruptcy Research Database, Mergent FISD©.

Covariates: financial ratios and market variables;
- Pre-processing: e.g., outlier removal by winsorization, removal of missing values.
Long-term issuer credit ratings assigned by S&P, Moody’s and Fitch for US companies excluding the financial and utilities sectors;
- S&P: AAA, AA, A, BBB, BB, B, CCC, CC
- Fitch: AAA, AA, A, BBB, BB, B, CCC, CC, C
- Moody’s: Aaa, Aa, A, Baa, Ba, B, Caa, Ca
- Sources: Compustat North America© Ratings File, Moody’s Default & Recovery Database©, Fitch Rating Services.

Failure indicator: binary indicator set to one on occurrence of bankruptcy filing under Chapter 7 or Chapter 11, or default rating by CRAs in the one year-window following the rating observation;
- Sources: UCLA-LoPucki Bankruptcy Research Database, Mergent FISD©.

Covariates: financial ratios and market variables;
- Pre-processing: e.g., outlier removal by winsorization, removal of missing values.

Period: 1999–2013
A first simple example

Model formula

```r
> formula <- MMO2(SPR, Fitch) ~ 0 + R3 + R12 + R18 + R20 + R23 + R34 + R35 +
+    RSIZE + BETA + SIGMA + MB + fyear
```

Constraints on thresholds

```r
> threshold.constraints <- c(1, 1)
```

Function call

```r
> res_SPR_Fitch <- mvord(formula,
+    data = data_ordinal,
+    threshold.constraints = c(1, 1),
+    link = mvlogit(),
+    error.structure = cor_general(~1))
```
A first simple example - Standardized regression coefficients

EBIT/interest expenses

debt/assets

LT debt/LT capital

retained earnings/assets

return on capital

R&D/assets

capital expenditures/assets

RSIZE

BETA

SIGMA

market to book

Examples
A first simple example - Year intercepts

\[ \exp(\alpha_t) = \frac{P(Y_{t,j} > r)/P(Y_{t,j} \leq r)}{P(Y_{1999,j} > r)/P(Y_{1999,j} \leq r)}, \quad j \in \{S&;P, Fitch\} \]
A joint model of credit ratings

Model formula

```r
> formula <- MMO2(SPR, Moodys, Fitch) ~ 0 + R3 + R7 + R12 + R18 + R20 + R23 + R24 +
+ R34 + R35 + RSIZE + BETA + SIGMA + MB
```

Constraints on regression coefficients

```r
> coef.constraints <- c(1, 1, 1)
```

Function call

```r
> res_SPR_Moodys_Fitch <- mvord(formula,
+   data = data_ordinal,
+   coef.constraints = c(1, 1, 1),
+   link = mvlogit(),
+   error.structure = cor_general(~1))
```
A joint model of credit ratings - Threshold coefficients

![Diagram showing credit rating thresholds for Fitch, Moody's, and S&P]

- Fitch
- Moody's
- S&P

Examples:

Ca
CCC/C_Caa
B
BB_Ba
BBB_Baa
A
AA_Aa
AAA_Aaa

Latent variable

-20 -18 -16 -14 -12 -10 -8 -6 -4 -2 0 1 2
We assume that S&P ($S$), Moody’s ($M$) and Fitch ($F$) provide ratings on an ordinal scale based on a latent process:

\[
\tilde{S}_i = x_i^\top \beta_S + \epsilon_i, S, \\
\tilde{M}_i = x_i^\top \beta_M + \epsilon_i, M, \\
\tilde{F}_i = x_i^\top \beta_F + \epsilon_i, F,
\]

where $\beta_S$, $\beta_M$ and $\beta_F$ are vectors of coefficients and $\epsilon_i, .$ are error terms.

For a binary default or failure indicator (labeled by $D$) we assume:

\[
\tilde{D}_i = x_i^\top \beta_D + \epsilon_i, D,
\]

where $\epsilon_i, D$ is a failure indicator specific error term.

For the errors we assume $[\epsilon_{ij}]_{j \in \{S, M, F, D\}} \sim F_{i, q_i}(0, R_i)$. 

A joint model of credit ratings and defaults

Model formula

```r
> formula <- MMO2(SPR, Moodys, Fitch, failInd) ~ 0 + R20 + R23 + R34 + SIGMA + 
  + BETA + R1 + R13 + R18 + lAT + MB + R1d + R5 + R17M + R22M + R27a + R29 + R35a
```

Constraints on coefficients

```r
> coef.constraints <- cbind(c(1,2,3,NA), c(1,2,3,NA), c(1,2,3,NA), c(1,2,3,4), 
  + c(1,2,3,NA), c(1,2,3,NA), c(1,2,3,NA), c(1,2,3,NA), c(1,2,3,4), c(1,2,3,NA), 
  + c(NA,NA,NA,1), c(NA,NA,NA,1), c(NA,NA,NA,1), c(NA,NA,NA,1), c(NA,NA,NA,1), 
  + c(NA,NA,NA,1), c(NA,NA,NA,1))
```

Function call

```r
> res_joint <- mvord(formula, data = data_ordinal, link = mvlogit(), 
  + weights = "weights3raters", coef.constraints = coef.constraints, 
  + error.structure = cor_general(~1))
```
Correlation

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P</th>
<th>Moody's</th>
<th>Fitch</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
<td>0.34</td>
<td>0.28</td>
</tr>
<tr>
<td>(0.019)</td>
<td>(0.017)</td>
<td>(0.027)</td>
<td>(0.132)</td>
<td>(0.119)</td>
</tr>
</tbody>
</table>

Examples
Evaluating out-of-sample predictive performance

The proposed model allows to predict PDs conditional on the observed ratings from the CRAs:

\[ P(D_i = 1 | S_i = r_i S, M_i = r_i M, F_i = r_i F) = \frac{P(D_i = 1, S_i = r_i S, M_i = r_i M, F_i = r_i F)}{P(S_i = r_i S, M_i = r_i M, F_i = r_i F)}, \]

where \( S_i, M_i \) and \( F_i \) denote the rating observations and \( D_i \) is the default indicator.

Cumulative Accuracy Profile

\[ \text{AR} = \frac{B}{A} = 0.901 \]

\[ \text{AUC B} = 0.444 \]

\[ \text{AUC A} = 0.492 \]
The underlying latent process of the proposed model is assumed to have the following form:

\[ \tilde{Y}_{it} = \mathbf{x}_{it}^\top \beta_t + \epsilon_{it}, \]

where

- \( \beta_t \) is a time-specific regression coefficient,
- \( \epsilon_{it} \) is an error term with an autocorrelation structure of order one (AR(1)):

\[ \epsilon_{it} = \rho \epsilon_{i(t-1)} + \sqrt{1 - \rho^2} \eta_{it}, \]

\[ \eta_{it} \sim \mathcal{N}(0, 1). \]
Longitudinal model of S&P ratings

Model formula

```r
> formula <- MMO(SPR, gvkey, fyear) ~ 0 + R3 + R9 + R12 + R18 + R20 + R23 +
+     R24 + R34 + R35 + RSIZE + BETA + SIGMA + MB
```

Threshold constraints

```r
> threshold.constraints <- rep(1, nlevels(data_ordinal$fyear))
```

Function call

```r
> res_ar1 <- mvord(formula, data = data_ordinal, link = mvprobit(),
+                   weights = "weights_SPR", threshold.constraints = threshold.constraints,
+                   error.structure = cor_ar1(~1))
```
Time varying coefficients (I)

EBIT/interest expenses

net PPE/assets

debt/assets

long term debt/long term capital

retained earnings/assets

return on capital

Examples
Time varying coefficients (II)

- EBITDA/sales
- R&D/assets
- capital expenditures/assets
- RSIZE
- BETA
- SIGMA
- market to book

Examples
The R package **mvord** can be used to:

- create multivariate shadow ratings,
- gain insights into the rating behavior,
- investigate the heterogeneity among CRAs,
- measure association between ratings and failures,
- build a joint model of failures and credit ratings which allows to perform inference about the relationship between these outcomes,
- provide interesting insights from the joint distribution, i.e., conditional probabilities can be computed.
Flexible modeling framework for multivariate ordinal regression models with:
- outcome-specific threshold coefficients,
- outcome-specific regression coefficients,
- constraints on threshold and regression parameters,
- different error structures and
- two multivariate link functions.

Package `mvord` is available on CRAN (Version 0.3.0).

A comprehensive package vignette is available.


References II


Thank you for your attention!

Rainer Hirk
rhirk@wu.ac.at

Institute for Statistics and Mathematics
WU Vienna University of Economics and Business
Welthandelsplatz 1
1020 Vienna