Dependent Hidden Markov Model of Credit Quality

Małgorzata Korolkiewicz
Centre for Industrial and Applied Mathematics
School of Mathematics & Statistics
University of South Australia
Outline

- Hidden Markov model of credit quality
- Dependent dynamics
- Recursive filter
- Parameter estimates
- Implementation example
Key references


1. Hidden Markov model of credit quality

- Under a probability measure $P$:
  \[
  X_{k+1} = AX_k + V_{k+1} \quad \text{(signal equation, “true” credit quality)}
  \]
  \[
  Y_{k+1} = CX_k + W_{k+1} \quad \text{(observation equation, posted rating)}
  \]
- $A$ and $C$ are matrices of transition probabilities whose entries satisfy
  \[
  \sum_{j=1}^{N} a_{ji} = 1, \quad a_{ji} \geq 0, \quad \sum_{j=1}^{M} c_{ji} = 1, \quad c_{ji} \geq 0.
  \]
- $V_k$ and $W_k$ are martingale increments satisfying
  \[
  E[V_{k+1} | \mathcal{F}_k] = 0
  \]
  \[
  E[W_{k+1} | \mathcal{G}_k] = 0.
  \]
2. Dependent dynamics

- Suppose noise terms $V_k$ and $W_k$ are not independent.
- Suppose that the joint distribution of $Y_k$ and $X_k$ is given by
  \[ Y_{k+1}X'_{k+1} = SX_k + \Gamma_{k+1}, \quad k = 0, 1, \ldots. \]
- $S = (s_{rji})$ denotes a $(M \times N) \times N$ matrix mapping $\mathbb{R}^N$ into $\mathbb{R}^M \times \mathbb{R}^N$, with $s_{rji} = P(Y_{k+1} = f_r, X_{k+1} = e_j \mid X_k = e_i)$, $1 \leq r \leq M, 1 \leq i, j \leq N$.
- $\Gamma_{k+1}$ is a $\{G_k\}$-martingale increment with $E[\Gamma_{k+1} \mid G_k] = 0$.
- Write $\gamma_{rji} = P(Y_{k+1} = f_r \mid X_{k+1} = e_j, X_k = e_i) = \frac{s_{rji}}{a_{ji}}$ and let $\tilde{C}$ be the $M \times (N \times N)$ matrix $(\gamma_{rji})$, $1 \leq r \leq M, 1 \leq i, j \leq N$. 

CEF 2009 5 17 July 2009
2.1 Dependent Hidden Markov model

- Under a probability measure $P$:
  
  $$X_{k+1} = AX_k + V_{k+1}$$
  $$Y_{k+1} = \tilde{C}(X_{k+1}X'_k) + \tilde{W}_{k+1}.$$  

- $A$ and $\tilde{C}$ are matrices of transition probabilities whose entries satisfy $\sum_{j=1}^N a_{ji} = 1$, $\sum_{r=1}^M \gamma_{rji} = 1$.

- $V_k$ and $\tilde{W}_k$ are martingale increments satisfying
  
  $$E[V_{k+1}|\mathcal{F}_k] = 0$$
  $$E[\tilde{W}_{k+1}|\mathcal{G}_k] = 0.$$
2.2. Remarks

- In this setting, the current credit rating contains information about both current and previous credit quality.

- This allows for the situation where a rating does not immediately reflect all available information about credit quality.

- Probabilities $\gamma_{rji}$ provide the distribution of the next period’s credit rating given both current and next period’s credit quality.
2.3. Test for independence

- If the noise terms in the state $X$ and observation $Y$ are independent, we have

$$P(Y_{k+1} = f_r, X_{k+1} = e_j | X_k = e_i) =$$

$$= P(Y_{k+1} = f_r | X_k = e_i) P(X_{k+1} = e_j | X_k = e_i).$$

- A test for independence is to check whether parameter estimates satisfy

$$\hat{s}_{r\cdot j} = \hat{c}_{r\cdot i} \hat{a}_{j\cdot i}.$$
3. Recursive filter

• Suppose we observe $Y_0, \ldots, Y_k$, and we wish to estimate $X_0, \ldots, X_k$.

• Under some probability measure $\overline{P}$ on $(\Omega, \mathcal{F})$, $\{Y_k\}$ is a sequence of i.i.d. uniform variables, and $X$ is Markov chain independent of $Y$, with state space $S = \{e_1, \ldots, e_N\}$ and transition matrix $A = (a_{ji})$.

• Define $\overline{\lambda}_l = M \sum_{j=1}^{M} \langle \tilde{C}(X_l X'_{l-1}), f_j \rangle \langle Y_l, f_j \rangle$ and $\overline{\Lambda}_k = \prod_{l=1}^{k} \overline{\lambda}_l$, and a new probability measure $P$ by putting $\frac{dP}{d\overline{P}}|_{g_k} = \overline{\Lambda}_k$. 
3.1. Best (mean-square) estimate

- The best (mean-square) estimate of $X_k$ given $\mathcal{Y}_k = \sigma\{Y_0, \ldots, Y_k\}$ is $E[X_k | \mathcal{Y}_k] \in \mathbb{R}^N$.

- However, $\tilde{P}$ is a much easier measure under which to work.

- Using Bayes’ Theorem we have

$$E[X_k | \mathcal{Y}_k] = \frac{E[\Lambda_k X_k | \mathcal{Y}_k]}{E[\Lambda_k | \mathcal{Y}_k]} := \frac{\tilde{q}_k}{\langle \tilde{q}_k, 1 \rangle}.$$ 

- A recursive formula for $\tilde{q}_{k+1}$ is given by

$$\tilde{q}_{k+1} = MY'_{k+1} S \tilde{q}_k.$$
4. Parameter estimation

To estimate parameters of the model, matrices $A$, $C$ and $S$, we need estimates of the following processes:

- $J_{ij}^k = \sum_{n=1}^{k} \langle X_{n-1}, e_i \rangle \langle X_n, e_j \rangle, 1 \leq i, j \leq N,$
- $O_i^k = \sum_{n=1}^{k} \langle X_{n-1}, e_i \rangle, 1 \leq i \leq N,$
- $T_{ir}^k = \sum_{n=1}^{k} \langle X_{n-1}, e_i \rangle \langle Y_n, f_r \rangle, 1 \leq i \leq N, 1 \leq r \leq M,$
- $L_{ijr}^k = \sum_{n=1}^{k} \langle X_{n-1}, e_i \rangle \langle X_n, e_j \rangle \langle Y_n, f_r \rangle, 1 \leq r \leq M, 1 \leq i, j \leq N.$
4.1. Best (mean-square) estimates

- As for the state of the Markov chain, for $H = \mathcal{J}, \mathcal{O}, \mathcal{T}, \mathcal{L}$ we wish to estimate

$$E[H_k | \mathcal{Y}_k] = \frac{E[\Lambda_k H_k | \mathcal{Y}_k]}{E[\Lambda_k | \mathcal{Y}_k]} := \frac{\sigma(H)_k}{\langle \tilde{q}_k, 1 \rangle}.$$  

- However, closed form recursions can only be obtained for

$$\sigma(HX)_k = E[\Lambda_k H_k X_k | \mathcal{Y}_k].$$

- Note that $\sigma(H)_k = \langle \sigma(HX)_k, 1 \rangle$. 
4.2. Recursive formulae

- $\sigma(J_{ij} X)_{k+1} =
  MY'_{k+1} S\sigma(J_{ij} X)_k + \langle \tilde{q}_k, e_i \rangle \left( M \sum_{r=1}^M s_{rji} \langle Y_{k+1}, f_r \rangle \right) e_j,$

- $\sigma(O_i X)_{k+1} =
  MY'_{k+1} S\sigma(O_i X)_k + \langle \tilde{q}_k, e_i \rangle \sum_{j=1}^M \left( M \sum_{r=1}^M s_{rji} \langle Y_{k+1}, f_r \rangle \right) e_j,$

- $\sigma(T_{ir} X)_{k+1} =
  MY'_{k+1} S\sigma(T_{ir} X)_k + M \langle \tilde{q}_k, e_i \rangle \left( \sum_{j=1}^N s_{rji} e_j \right) \langle Y_{k+1}, f_r \rangle,$

- $\sigma(L_{ijr} X)_{k+1} = MY'_{k+1} S\sigma(L_{ijr} X)_k + \langle \tilde{q}_k, e_i \rangle M s_{rji} \langle Y_{k+1}, f_r \rangle e_j.$
4.3. Parameter estimates

- Given the observations up to time \( k \), \( \{ Y_0, Y_1, \ldots, Y_k \} \), and given the parameter set \( \theta = \{ a_{ji}; c_{ji} \} \), the EM estimates \( \hat{a}_{ji} \) and \( \hat{c}_{ji} \) are given by
  \[
  \hat{a}_{ji} = \frac{\sigma(T^{ij})_k}{\sigma(\mathcal{O}_i)_k}
  \]
  and
  \[
  \hat{c}_{ji} = \frac{\sigma(T^{ij})_k}{\sigma(\mathcal{O}_i)_k}.
  \]

- Given the observations up to time \( k \), \( \{ Y_0, Y_1, \ldots, Y_k \} \), and given the parameter set \( \theta = \{ a_{ji}; c_{ji}; s_{r_{ji}} \} \), the EM estimates \( \hat{s}_{r_{ji}} \) are then given by
  \[
  \hat{s}_{r_{ji}} = \frac{\sigma(L^{ijr})_k}{\sigma(\mathcal{O}_i)_k}.
  \]
5. Implementation example

- Standard & Poor’s COMPUSTAT rating histories.
- Annual ratings for mainly large U.S. and Canadian corporate institutions: industrials, utilities, insurance companies, banks and other financial institutions and real estate companies.
- Every year each of the rated obligors is assigned to one of the Standard and Poor’s 7 rating categories, ranging from AAA (highest rating) to CCC (lowest rating) as well as D (payment in default) and the NR (not rated) state.
- Given the fairly large number of parameters to be estimated compared to the number of rating transitions in the dataset, all firms in the dataset were reclassified as IG (investment grade), SG (speculative grade), D or NR.
## Estimated matrix A

<table>
<thead>
<tr>
<th></th>
<th>IG</th>
<th>SG</th>
<th>D</th>
<th>NR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IG</td>
<td>0.408</td>
<td>0.018</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>SG</td>
<td>0.068</td>
<td>0.249</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>D</td>
<td>0.000</td>
<td>0.017</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>NR</td>
<td>0.524</td>
<td>0.715</td>
<td>0.000</td>
<td>0.999</td>
</tr>
</tbody>
</table>

## Estimated matrix C

<table>
<thead>
<tr>
<th></th>
<th>IG</th>
<th>SG</th>
<th>D</th>
<th>NR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IG</td>
<td>0.126</td>
<td>0.038</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>SG</td>
<td>0.094</td>
<td>0.118</td>
<td>0.010</td>
<td>0.000</td>
</tr>
<tr>
<td>D</td>
<td>0.001</td>
<td>0.004</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>NR</td>
<td>0.780</td>
<td>0.840</td>
<td>0.990</td>
<td>1.000</td>
</tr>
</tbody>
</table>
5.1. Some observations

- Investment-grade firms generally hold on to their status.
- Speculative-grade firms tend to maintain their status or disappear from the data set.
- Transition to NR status is more likely for speculative-grade than for investment-grade obligors.
- Rating agencies may be reluctant to upgrade (downgrade) firms to (from) investment-grade status.
- Upon checking whether estimated parameters satisfied \( \hat{s}_{rji} = \hat{c}_{ri} \hat{a}_{ji} \), no significant departures from independence were found.
- However, longer rating histories may be necessary to verify these findings.