

Modelowanie ryzyka kredytowego

Wykorzystanie panelowej regresji liniowej

Stanisław Ochotny
Credit Methodology – FI / Corporates

28 września 2017



Podjęcie klasyczne

Do budowy modelu służy zazwyczaj historia niespłacalności

- Modele ryzyka kredytowego dla portfela powstają najczęściej w oparciu o historyczne zdarzenia niespłacalności
- Do modelowania prawdopodobieństwa niespłacalności możemy wykorzystać regresję logistyczną:

$$\text{logit}(p_i) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_K x_{K,i}$$

- W przypadku klienta korporacyjnego, zmiennymi objaśniającymi są najczęściej dane opisujące strukturę bilansu oraz rachunek zysków i strat
- Model zakłada, że dostępna jest duża i wiarygodna historia niespłacalności danego portfela

Rating odzwierciedlający

Jeżeli dane historyczne nie są dostępne, rozważamy inne opcje

- Rating odzwierciedlający (shadow rating) w założeniu ma replikować zewnętrzne miary oceny ryzyka kredytowego i być stosowalny także do firm, dla których zewnętrzna ocena ryzyka nie jest dostępna
- Poziom zgodności modelu odzwierciedlającego mierzy się korelacją rank (τ Kendalla, D Somersa)
- Zewnętrzne oceny ryzyka kredytowego:
 - Oceny agencji ratingowych,
 - PD implikowane z kredytowych instrumentów pochodnych (CDS)
 - Expected Default Frequency,

Odległość do niespłacalności

Niespłacalność modelujemy używając zmiennej ukrytej

- Zewnętrzne miary ryzyka kredytowego, jeżeli mają postać prawdopodobieństwa niespłacalności, transformujemy funkcją: $f: (0,1) \rightarrow (0, \infty)$
- Tak uzyskaną miarę określamy jako odległość do niespłacalności (distance to default) i zakładamy, że zjawisko niespłacalności zachodzi, gdy odległość ta osiągnie zero
- Struktura danych:
 - Obserwujemy N podmiotów, dla których gromadzimy zewnętrzne miary ryzyka kredytowego i ich sprawozdania finansowe przez okres T lat
 - Obserwacje tworzą grupy w zależności od podmiotu, którego dotyczą, co sugeruje panelową strukturę danych
 - Ponieważ dla niektórych podmiotów obserwacje nie są dostępne przez wszystkie T lat, mamy do czynienia z panelem niezbilansowanym

Efekty nieobserwowalne

Zmienne, które używamy w modelu, mogą nie wyjaśniać wszystkiego

- Odległość do niespłacalności może być konsekwentnie powyżej lub poniżej wartości, która wynika z zastosowanych zmiennych objaśniających
- Budując model ilościowy ryzyka kredytowego, trudno uwzględnić:
 - Rozpoznawalność marki i skuteczność marketingu,
 - Jakość zarządzania spółką,
 - Strukturę organizacyjną przedsiębiorstwa,
 - Kulturę i etykę pracy w firmie.
- Dlatego modelując dane panelowe, uwzględniamy dodatkowe, nieobserwowalne efekty indywidualne
- Równanie modelu przyjmuje postać:

$$y_{i,t} = \mathbf{x}_{i,t}\boldsymbol{\beta} + c_i + u_{i,t}$$

gdzie $\mathbf{x}_{i,t}$ jest wektorem zmiennych objaśniających, c_i efektem indywidualnym, natomiast $u_{i,t}$ jest błędem losowym

Zwykła regresja liniowa

Traktowanie wszystkich obserwacji jak niezależnych prowadzi do złych wyników

- Dokonując regresji na wszystkich dostępnych obserwacjach, jak gdyby były danymi przekrojowymi, prowadzi do nieefektywnych i obciążonych wyników
- Przed podjęciem decyzji, czy zastosować klasyczną regresję liniową, czy regresję panelową, stosujemy testy na istnienie efektów indywidualnych:
 - Test Breuscha-Pagana (1980),
 - Test Hondy (1985),
 - Test Kinga-Wu (1997),
 - Test Gourieroux, Holly'ego i Monforta (1982)
- Wszystkie testy zaimplementowane są w pakiecie plm i wywoływane funkcją `plmtest` z odpowiednimi parametrami

Panelowa regresja liniowa

Dwa podejścia do efektów indywidualnych

Równanie panelowej regresji liniowej:

$$y_{i,t} = x_{i,t}\beta + c_i + u_{i,t}$$

- Istnieją dwa zasadnicze podejścia do efektów indywidualnych:
 - Traktować je jako zmienne losowe (random effects),
 - Lub jako stałe, parametry które podlegają estymacji (fixed effects).
- Oba modele estymujemy funkcją plm z pakietu plm, z parametrami `method = "random"` lub `method = "within"`.
- To, które podejście możemy stosować, zależy od relacji między efektami indywidualnymi i zmiennymi objaśniającymi:
 - Model efektów losowych zakłada brak korelacji między efektami indywidualnymi i zmiennymi objaśniającymi,
 - Model efektów stałych dopuszcza dowolną korelację między efektami indywidualnymi i zmiennymi objaśniającymi.

Test Hausmana

Metoda sprawdzenia istnienia korelacji

- Jeżeli nie ma korelacji między efektami indywidualnymi, a zmiennymi objaśnianymi, model efektów losowych jest bardziej efektywny
- W przypadku istnienia korelacji między efektami indywidualnymi i zmiennymi objaśniającymi, estymator efektów losowych nie jest zgodny
- Test Hausmana, zaproponowany w 1978, pozwala na testowanie hipotezy o zgodności modelu z efektami losowymi
- Test jest zaimplementowany w pakiecie plm i wywoływany funkcją `phtest`, gdzie parametrami są dwa modele – jeden z efektami losowymi, drugi z efektami stałymi

Pozostałe problemy

Rzeczy wymagające więcej uwagi

- Używając modelu ze stałymi efektami indywidualnymi, trudno ekstrapolować model poza próbę użytą do estymacji
- Model panelowej regresji liniowej zakłada brak autokorelacji błędów idiosynkratycznych (choć jej występowanie nie wpływa na niezgodność estymatora)
- Testowanie występowania autokorelacji możemy dokonać używając funkcji `pbsytest`, `pbltest`, `pbgtest`.
- W przypadku istnienia autokorelacji błędów, potrzebujemy odpornego (robust) estymatora wariancji dla otrzymanych parametrów modelu
- Odporne estymacje wariancji możemy uzyskać, testując czy otrzymane parametry są znaczące (`coeftest`), używając funkcji `vcovHC` z pakietu `lmtest`.

Literatura

Podręczniki akademickie, regulacje bankowe, dokumentacja R

- Wooldridge, J., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2010.
- Basel Committee on Banking Supervision, *Studies on the Validation of Internal Rating Systems*, Bank for International Settlements, Working Paper No. 14.
- Croissant Y., Millo G., *Panel Data Econometrics in R: The plm Package*

Informacje kontaktowe

Stanisław Ochotny
Credit Methodology – FI / Corporates
stanislaw.ochotny@ubs.com

www.ubs.com

