

Aalto-yliopisto

Mat-2.4177 Operaatiotutkimuksen projektityöseminaari

Kevät 2012

Luottoluokitusten siirtymätodennäköisyyksien estimointi ja kalibrointi

Väli­raportti

12.4.2012

Janne Kunnas (projektipäällikkö)
Tuomas Nikoskinen
Joonas Parjanne
Raul Kleinberg
Mikko Kotilainen

1. Projektin eteneminen

Projekti on edennyt hyvin suunnitelman mukaan. Erona alkuperäiseen suunnitelmaan on se, että kirjallisuuskatsaus ei ole vielä valmis, mutta sovellettavien mallien osalta ollaan aikataulusta edellä. Transitiomatriisin kalibrointi on suoritettu suhdannemallilla ja tulosten analysointi valmistuneen aikataulusta kaksi viikkoa etujassa. Tulosten analysoinnin osalta on varauduttava suunniteltua suurempaan työmäärään, koska käytettyjen mallien, suhdannemalli ja sekoitemalli, tulosten vertailu voi olla hankalaa, perustelut esitetty kappaleessa 4 Riskit. Suurin osa jäljellä olevista tehtävistä on loppuraportin kirjoittamista.

1.1. Työnjako

Projektiryhmä on jaettu kahteen eri ryhmään. Ensimmäisessä ryhmässä on Janne, Joonas ja Mikko, jotka ovat kehittäneet suhdannemallia. Suhdannemalli on tässä vaiheessa projektia validointivaiheessa ja jäljellä on mallin kehittämisen osalta enää lopullisen transitiomatriisin vertailu sekoitemallin tuottaman transitiomatriisin kanssa. Sekoitemalliryhmään kuuluu Tuomas ja Raul. Loppuraportin vastualueet ja arvioidut työmäärät on esitelty taulukossa 1.

Taulukko 1. Projektin tehtävien jako loppuraportin osalta.

TEHTÄVÄ	TEKIJÄT	TYÖMÄÄRÄ (sivua)
Johdanto	Raul	1
Ongelman määrittely	Tuomas	2
Kirjallisuuskatsaus	Mikko, Janne, Joonas	3
Menetelmät	Tuomas, Raul, Janne, Joonas	4
Mallin valinta	Joonas	2
Datan analysointi	Tuomas, Raul	5
Estimointi suhdannemallilla	Janne, Joonas	4
Estimointi sekoitemallilla	Tuomas, Raul	4
Tulokset	Joonas	2
Pohdinnat	Kaikki	2

1.2. Aikataulu

Projektin päivitetty aikataulu on esitetty taulukossa 2. Taulukon oikeassa reunassa oleva tyhjä viikko on jätetty pelivaraa varten. Tarkoitus olisi saada projekti valmiiksi viikolla 18. Loppuraportin palautuksen takaraja on viikolla 19. Suorittamattomia tehtäviä ovat vielä kirjallisuuskatsaus, tulosten

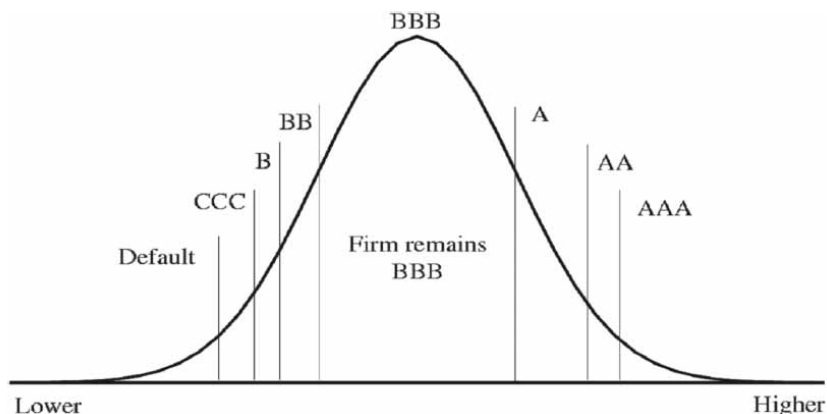
analysointi sekä loppuraportin kirjoittaminen. Näistä tehtävistä kaikki on aloitettu jo sekä vastualueet jaettu jokaiselle ryhmän jäsenelle.

Taulukko 2. Projektin aikataulu ja suoritettavat tehtävät.

TEHTÄVÄ		HELMIKUU	MAALISKUU	HUHTIKUU	TOUKO
		SUUNNITELMA	VÄLIRAPORTTI	LOPPURAPORTTI	
Työn suunnittelu	Suoritettu	■			
Lähdemateriaalin keräys	Suoritettu	■			
Lähdemateriaalin läpikäyminen	Suoritettu	■			
Projektsuunnitelma	Suoritettu	■			
Mallivaihtoehtojen tutkiminen	Suoritettu	■			
Mallien valinta	Suoritettu		■		
Kirjallisuuskatsaus			■		
Transitiomatriisin estimointi	Suoritettu		■		
Mallien validointi	Suoritettu		■		
Väliraportti	Suoritettu			■	
Transitiomatriisin kalibrointi	Suoritettu			■	
Tulosten analysointi				■	
Loppuraportti				■	

2. Suhdannemalli

Olemme toteuttaneet Matlabilla toimintasuunnitelman mukaisesti mallin, joka huomioi suhdannevaihteluiden vaikutuksen luottoluokitusten transitiotodennäköisyyksien estimoinnissa. Mallissa liitetään kuhunkin luottoluokitukseen i standardinormaalijakaumaa noudattava satunnaismuuttuja, luottopistemäärä seuraavana vuonna (kuva 1). Tämän mielekkyyttä voidaan tarkastella sovittamalla normaalijakauma transitiomatriisin riviin i . Jakauma voidaan osittaa eri luottoluokituksiin j siirtymistä vastaaviin osiin, joiden todennäköisyydet ovat transitiomatriisin alkio p_{ij} .



Kuva 1. Tulevan luottopistemäärän normaalijakauma ja eri luokkiin siirtymistä vastaavat osat luokalle BBB. Usein suhdannevaihtelujen vaikutuksen arvioimiseksi

tarkastellaan keskinkertaisia luottoluokkia, joihin suhdannevaihtelut usein vaikuttavat selvemmin kuin korkeisiin luottoluokkiin. Myös normaalijakaumaoletus toimii parhaiten keskinkertaisille luottoluokille.

Suhdannealli kuvaa, kuinka maksuhäiriöön johtavan luottopistemäärän kynnyсарvo riippuu edeltävän ajanhetken makrotaloudellisista tilastoluvuista ja saa transitiotodennäköisyydet heilahtelemaan suhdannevaihteluiden mukana.

Olemme yllä kuvattua mallia käyttäen laskeneet ennusteen vuoden 2012 transitiotodennäköisyysmatriisille. Käytimme regressiomallissa selittävinä muuttujina Tilastokeskuksen kuluttajahintaindeksi- ja työttömyysastelukuja, sillä näitä käyttäen saatiin regressiomallille korkein selitysaste.

Merkittävin haaste suhdannemalliin liittyen on vähäisellä datalla pärjääminen. Datan vähäisyyden vuoksi mallimme selitysaste uhkaa jäädä alhaiseksi. Yritämme selvittää mahdollisuuksia saada toimeksiantajalta lisää dataa tai keksiä järkeviä tapoja hyödyntää ulkopuolista dataa kuten Tilastokeskuksen konkurssidataa 1994-2011.

Mallin verifioimiseksi voidaan esimerkiksi ennustaa 2011 transitiomatriisi sitä aiemmasta datasta huomioiden suhdannevaihtelu ja verrata ennustetta toteutuneeseen. Tällöin datan vähäisyys kuitenkin korostuu entisestään. Mallia tulisi lisäksi verrata myös kehittämäämme sekoitetemalliin. Lisäksi tulisi suorittaa herkkyysanalyysi suhdannemallin estimoinnissa käytetyille parametreille.

3. Sovitemalli

3.2. Mallin konstruoinnin eteneminen

Projektisuunnitelmassa sovitemallin lähtökohdaksi asetettiin todennäköisyysjakauman sovittaminen suoraan aiheenasettajan aineistosta laskettujen luottoluokitusten siirtymäfrekvensseihin. Kuvassa 2 on esitetty pääpiirteittäin kuinka sovitemallin konstruointi on edennyt lähtötilanteesta sekoitemallin löytämiseen.



Kuva 2. Sovitemallin konstruoinnin eteneminen.

Ensimmäinen kokeiluluontoinen idea oli sovittaa yhteisjakauma siirtymäfrekvensseihin kokeillen sekä parametrisia että ei-parametrisia malleja. Ongelmaksi tässä lähestymistavassa muodostui havaintojen vähyys sekä niiden jakautumien pieneen määrään diskreettejä pisteitä (luottoluokitusluokat).

Yhteisjakaumaan pohjautuvasta lähestymistavasta päätettiin luopua siihen liittyvien ongelmien ja haasteiden vuoksi. Luonnollinen seuraavana askel oli ehdollistaa siirtymätodennäköisyysjakaumat luottoluokituskohtaisiksi, jolloin kunkin luokan siirtymille estimoidaan oma yksiulotteinen todennäköisyysjakauma. Näiden jakaumien haluttiin muistuttavan kvalitatiivisilta ominaisuuksiltaan tarkasteltavaa aineistoa sekä olevan samasta jakaumaperheestä.

Visuaalisen tarkastelun perusteella todettiin, että normaali- ja t-jakauma täyttävät sovitettaville jakaumille asetetut kriteerit, jos tavanomaisten koko reaaliakselin yli määritettyjen jakauman sijaan käytetään nk. typistettyä jakaumaa, joka on rajoitettu annetulle intervallille (jolla luottoluokat on määritelty). Lisäksi, koska molemmat sovitettavat jakaumat ovat jatkuvia ja havainnot diskreettejä, päätettiin estimoinnin robustisuuden vuoksi hajauttaa havaintoja.

Typistettyjen jakaumien avulla saatiin estimoitua varsin järkevänolaisia siirtymäjakaumia. Estimaateissa oli kauttaaltaan kuitenkin ongelmana, että todennäköisyys, jolla siirtymää ei tapahtunut ("pysytään samassa luokassa"), oli aivan liian pieni. Tämä ilmiö selittyy ennen kaikkea sillä, että aineisto koostuu havainnoista, jotka on kerätty melko volatiilin periodin aikana, jolloin havainnot luokituksen muuttumisesta on suhteessa enemmän kuin havaintoja sen pysymisestä samana.

Huipukkuusongelman pienentämiseksi päätettiin hyödyntää ulkopuolisen luottoluokittajan (Fitch) tarjoamaa pitkältä aikaväliltä estimoitua keskimääräistä transitiomatriisia, jonka implikoimat luokittaiset transitiotodennäköisyys-jakaumat ovat huipukkaista. Hakemalla painotettu keskiarvo näistä huipukkaista jakaumista ja lasketuista estimaateista saadaan muodostettua sekamalli, joka yhdistää pitkän aikavälin keskiarvon ja viimeisimmät havainnot luottoluokitustransitioista lyhyeltä aikaväliltä.

3.3. Sekoitemalli

Sekoitemallin perusideana on mallintaa luottoluokitusten siirtymätodennäköisyyksiä ehdollisten luokituskohtaisten todennäköisyysjakaumien avulla. Kullekin luokalle i estimoidaan todennäköisyysjakauma, joka antaa siirtymätodennäköisyydet eri luottoluokkien välillä sekä todennäköisyyden siirtyä maksuhäiriöluokkaan. Nämä todennäköisyysjakaumat muodostetaan luokkakohtaisina sekoitemalleina, joissa aiheenasettajan aineistosta estimoidut jakaumat yhdistetään ulkopuolisen luottoluokittajan (Fitch) aineistosta estimoituihin vastaavan luokan jakaumiin.

Aiheenasettajan aineistosta luokkakohtaisiksi todennäköisyysjakaumiksi siirtymille estimoidaan typistetty normaali- ja t-jakauma. Näistä

jakaumatyypeistä valitaan lopulta validoinnin perusteella aineistoa paremmin kuvaava jakauma. Koska molemmat tarkasteltavat jakaumat ovat jatkuvia, hajautamme hieman luokkakohtaisia diskreettejä havaintoja, jotta havaintopisteitä saadaan enemmän estimointia varten. Hajautuksen vaikutusta estimointiin tullaan tarkastelemaan tarkemmin validoinnin yhteydessä. Lopullinen diskreetti transiiojakauma muodostetaan jatkuvasta jakaumasta kertymäfunktion avulla.

Ulkopuolisen luokittajan aineisto koostuu pelkästään frekventisesti lasketuista transiitodennäköisyyksistä luokkien välillä (ei sisällä informaatiota havaintomääristä). Tämän vuoksi diskreetteinä transiiojakaumina käytetään suoraan transiioatriisin rivejä.

3.4. Nykytila ja tulevaisuus

Sovitemallin lopulliseksi muodoksi on valittu yllä kuvattu sekoitemalli. Tämä malli on kokonaisuudessaan toteutettu ja työskentely painottuu sillä saatujen tulosten analysointiin, validointiin ja vertailuun suhdannemallin kanssa. Lopullisen mallinvalinnan kannalta on ratkaistavana vielä muutamia relevantteja kysymyksiä

- Valitaanko sekoitemallissa käytettäväksi jakaumaksi tyypistetty normaalijakauma vai tyypistetty t-jakauma?
- Saadaanko datan hajauttamisella lisäarvoa sekoitemalliin?
- Mitä painokerrointa käytetään sekoitemallin komponenttien yhdistämisessä ja miten painokerroin tulisi valita?

4. Riskit

Projektisuunnitelmassa arviomme, että datan laatu sekä estimointimallien sovellettavuus saattavat osoittautua suurella todennäköisyydellä haasteellisiksi. Suhdannevaihtelun osalta data on riittämätön. Tämän ongelman olemme ratkaisseet käyttämällä suhdanneindikaattorin implementoinnissa Tilastokeskuksen tarjoamaa makrodataa. Malleista bayesialainen osoittautui sen verran haasteelliseksi, että sen käsittelyssä tyydytään vain mallin esittelyyn, eikä sitä sovelleta aiheen asettajan aineistoon. Päivitetyt riskit on esitetty taulukossa 3. Uutena tunnistettuna riskinä on kahden eri mallien tuottamien transiioatriisien vertailtavuus. Koska estimointimallit, sekoitemalli ja suhdannemalli, ovat lähtökohtaisesti menetelmiltään hyvin erilaiset, on niiden tuottamien tulosten vertailu keskenään vaikeaa. Loppuraportissa tarkoituksena on vertailla mallien tuottamia transiioatriiseja yksinkertaisen Markov-mallin avulla estimoituun transiioatriisiin.

Taulukko 3. Projektin riskit.

RISKI	TODENNÄKÖISYYS	VAIKUTUS	TOIMENPITEET
Datan laatu	korkea	korkea	-Käytetään dataa myös ulkoisista lähteistä: Fitch, Moody's, Tilastokeskus
Työmäärä osoittautuu liian suureksi	kohtalainen	kohtalainen	-Vähennetään sovellettavien mallien määrää
Lopputulokset eivät tyydytä aiheenasettajaa	kohtalainen	kohtalainen	-Esitetään tulokset mahdollisimman yksinkertaisesti
Mallien soveltaminen osoittautuu ennakoitua haasteellisemmaksi	kohtalainen	korkea	-Jätetään varhaisessa vaiheessa tarkastelun ulkopuolelle haasteellisemmat mallit
Mallien tuottamien tulosten vertailu hankalaa	korkea	kohtalainen	