Aalto-yliopisto
Perustieteiden korkeakoulu
Teknillisen fysiikan ja matematiikan tutkinto-ohjelma

Menetelmä Markowitzin mallin parametrien estimointiin

kandidaatintyö
27.1.2016

Lauri Nyman

Aalto-yliopisto
Perustieteiden korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Tekniikan kandidaatintyön tiivistelmä

Tekijä Lauri Nyman	
Työn nimi Kandidaatintyö	
Koulutusohjelma Teknillisen fysiikan ja matematiikan tutkinto-ohjelma	
Pääaine Systeemitieteet	**Pääaineen koodi** F3010
Vastuuopettaja Prof. Harri Ehtamo	
Työn ohjaaja(t) TkT Eeva Vilkkumaa	
Päivämäärä 27.1.2016	**Sivumäärä** 15+3 · **Kieli** suomi

Tiivistelmä:

Sijoitustoiminnan tavoitteena on tyypillisesti löytää mahdollisimman pieniriskinen sijoitusportfolio mahdollisimman hyvällä tuotto-odotuksella. Jos sijoituskohteiden tuotto-odotutukset ja kovarianssit tunnetaan, voidaan tehokkaat portfoliot ratkaista Makrowitzin mallin avulla. Perinteisesti mallin parameteja eli tuotto-odotuksia ja kovariansseja on estimoitu historiatiedoista. Tulevaisuuden tuotot eivät kuitenkaan korreloi voimakkaasti historian kanssa ja instrumenttien keskinäiset korrelaatiot muuttuvat ajan mukaan. Sijoitusportfolion valinta on haastavaa ja siihen kaivataan uusia menetelmiä.

Tässä työssä kehitään uusi menetelmä Markowitzin portfoliomallin parametrien estimointiin. Tuotto-odotuksia estimoidaan faktorimallilla, eli erilaisten tuotto-odotuksiin vaikuttavien tekijöiden lineaarikombinaatiolla. Kovarianssien estimointiin kehitetty uusi menetelmä auttaa huomioimaan tuotto-odotusten estimoinnin epävarmuuden.

Estimaattien toimivuutta testataan eri maiden pörssiyhtiöistä koostuvilla ETF-rahastoilla, jotka pyrkivät jäljittelemään eri maiden pörssi-indeksejä. Testiaineiston perusteella instrumenttien tuotto-odotusten estimointiin käytetty faktorimalli vaikuttaa heikolta, mutta kovarianssien estimointimenetelmä toimi hyvin. Työssä käytetyillä estimaateilla saatiin keskimäärin paremmin tuottavat ETF-portfoliot kuin perinteisillä estimaateilla.

Avainsanat Tuotto-odotus, kovarianssin estimaatti, Markowitzin malli, tehokas rintama, tehokas portfolio, faktorimalli, tunnusluvut, Brownin liike.

Sisältö

1 Johdanto

Sijoitustoiminnalla on suuret markkinat ja sitä harjoitetaan maailmanlaajuisesti. Esimerkiksi vuonna 2013 New Yorkin pörssin osakkeita vaihdettiin päivässä keskimäärin 169 miljardin USA:n dollarin edestä vastaten suuruusluokaltaan Suomen bruttokansantuotetta.[1] Sijoitustoiminnan tavoitteena on tyypillisesti löytää mahdollisimman hyvin tuottava ja vähäriskinen kokoelma osakkeita eli sijoitusportfolio. Portfolion valinnan tukena on perinteisesti käytetty Harry Markowitzin 1950-luvulla kehittämää klassista portfoliohallinnan teoriaa.[2] Markowitzin mallin avulla saadaan määritettyä sijoitusosuudet eri sijoituskohteisiin siten, että sijoitusosuuksien määrittämillä portfolioilla on riskitasoonsa nähden suurin mahdollinen tuotto-odotus ja toisaalta odotettuun tuottoonsa nähden pienin mahdollinen riskitaso.[2]

Sijoituskohteiden odotetut tuotot ja tuottojen väliset kovarianssit oletetaan Markowitzin portfoliomallissa tunnetuiksi. Näiden parametrien arvoja ei kuitenkaan todellisuudessa tiedetä etukäteen, vaan ne pitää estimoida. Perinteisesti tulevaisuuden tuottoja ja kovariansseja estimoidaan historiasta, mutta tulevaisuuden tuotot eivät korreloi voimakkaasti historian kanssa ja instrumenttien keskinäiset korrelaatiot muuttuvat ajan mukaan. Myöhemmin on kehitetty uusia menetelmiä estimaattien parantamiseksi, kuten M-estimaatti kovarianssien estimointiin[10] ja faktorimalli tuotto-odotuksien estimointiin.[7] Erityisesti kovarianssien estimaattien ongelma on kuitenkin tuotto-odotusten estimointivirheen huomiotta jättäminen, joka on olennainen osa kokonaishajontaa.

Tässä työssä kehitetään uusi menetelmä Markowitzin mallin parametrien estimointiin. Odotettuja tuottoja estimoidaan faktorimallilla eli erilaisten tuottoon vaikuttavien tekijöiden (eli faktoreiden) lineaarikombinaatiolla. Kovarianssien estimoinnissa puolestaan otetaan huomioon tuotto-odotusten estimoinnin epävarmuus. Estimointimenetelmää testataan eri maiden pörssiyhtiöistä koostuvilla ETF-rahastoilla, jotka pyrkivät jäljittelemään eri maiden pörssi-indeksejä.

Työn rakenne on seuraavanlainen. Luvussa 2 esitellään aiemmin kehitettyjä menetelmiä Markowitzin mallin parametrien estimointiin. Luvussa 3 esitellään Makrowitzin malli ja tässä työssä käytetyt menetelmät tuottojen ja kovarianssien estimointiin. Luvussa 4 kuvataan testiaineisto ja tarkastellan miten malli olisi toiminut, jos sijoittaja olisi tehtyt sijoituspäätökset sen mu-

[1] https://www.nyse.com, New York Stock Exchange

kaan. Luvussa 5 on tehty johtopäätökset ja yhteenveto työstä.

2 Kirjallisuuskatsaus

Markowitzin mallin parametrien estimointi on tunnetusti hankalaa historiatietojen perusteella. Etenkin tuotto-odotusten estimaattien keskihajonta on tyypillisesti suuri.[3][6][7] Myös instrumenttien välinen korrelaatio muuttuu ajan mukaan; esimerkiksi sähkön ja öljyn hinnan korrelaatio muuttuu, kun sähköä aletaan tuottamaan uusiutuvilla energiamuodoilla. Talouden riippuvuussuhteet siis muuttuvat, eivätkä keskinäiskorrelaatiot käyttäydy samalla tavalla kovin monia vuosia tai vuosikymmeniä. Kovarianssien estimaatit on siis hyvä määrittää sopivalta aikajänteeltä.

Kirjallisuudessa estimaattien parantamiseksi on esitetty monenlaisia menetelmiä. Kan ja Zhou (2007) pyrkivät pienentämään estimaattien estimointivirheen vaikutusta valitun portfolioiden kokonaishajontaan yhdistämällä tehokkaita portfolioita siten, että odotettu tuotto pysyy kiinnitettynä.[5] Oliver ja Michael (2002) puolestaan pyrkivät pienentämään estimaattien hajontaa yhdistämällä erilaisia estimaatteja sopivilla painokertoimilla.[7] Lorenzo, Raman ja Tan (2007) muokkaavat portfolio-optimointiongelmaa siten, että instrumenttien tuotto-odotusten epävarmuus tulee huomioitua jo optimointivaiheessa.[8] Victor ja Francisco (2009) pyrkivät luomaan optimointifunktion, joka valitsee portfoliot, joissa instrumenttien osuudet eivät muutu niin voimakkaasti tuottojen ja kovarianssien mukaan.[9]

Eräs yleinen tapa tuottojen estimoinnissa on faktorimalli, jossa odotettuja tuottoja kuvataan erilaisten faktorien lineaarikombinaationa. Oliverin ja Michaelin (2002) mukaan faktorimallin muodostus on hankalaa, koska monen faktorin mallissa lähdeaineiston koko ja aikajänne ovat usein liian pieniä luotettavan mallin luomiseen.[7] Lisäksi faktorimallin toimivuuden validiointiin tarvitaan lisää aineistoa, joka on faktorimallin parametrien estimointiin käytetyn aineiston ulkopuolella.[7]

Edellä kuvatuilla menetelmillä on monesti päästy parempaan lopputulokseen kuin tavanomaisilla kovarianssien ja tuotto-odotusten estimaateilla.[7][8] Monet menetelmistä ovat kuitenkin monimutkaisia, eikä lopputulos ole parantunut oleellisesti. Parempia menetelmiä ja estimaatteja kaivataan edelleen.

3 Menetelmät

Tässä luvussa esitellään portfoliovalinnassa käytettävä Markowitzin malli sekä mallin parametrien (tuotto-odotuksen ja kovarianssit) estimointiin kehitettävät menetelmät.

3.1 Markowitzin malli

Markowitzin portfoliomallissa instrumentteja i ($i = 1, ..., N$), kuvataan niiden tuotto-odotuksilla $\mathbf{r} = (r_1, ..., r_N)^T$ ja kovarianssimatriisilla $\mathbf{\Sigma}$, jonka alkio Σ_{ij} kuvaa instrumenttien i ja j välistä kovarianssia. Sijoitusportfoliolla tarkoitetaan sijoitusosuusvektoria $\mathbf{w} = (w_1, ..., w_N)^T$, jolle $\sum_{i=1}^{N} w_i = 1$. Jos lisäksi on voimassa $w_i \geq 0 \; \forall \; i$, niin lyhyeksi myyntiä ei voi olla. Tavoitteena on löytää sijoitusportfolio $\mathbf{w}$, jolla on tuotto-odotukseensa nähden pienin mahdollinen varianssi $\mathbf{w^T \Sigma w}$. Tämän tyyppisiä portfolioita kutsutaan tehokkaiksi. Olkoon portfolion tuotto-odotus $R_0 = \mathbf{w^T r}$ kiinnitetty. Nyt tehokas portfolio voidaan määrittää ratkaisemalla seuraava optimointitehtävä:[1]

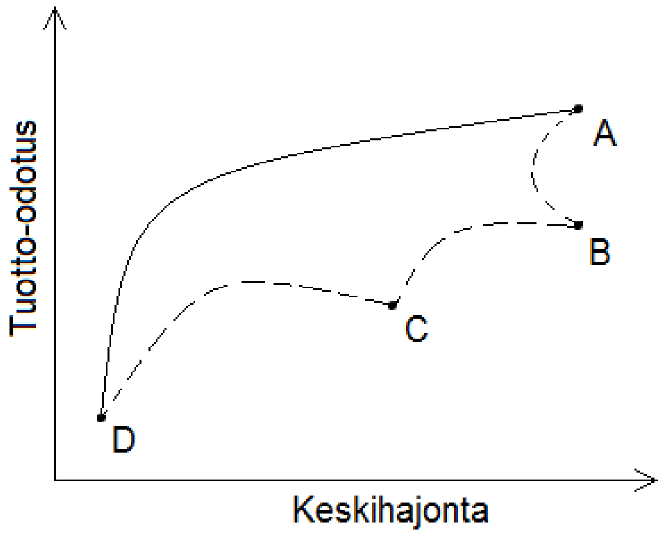
$$\min_{\mathbf{w}} \sigma^2 = \min_{\mathbf{w}} \left(\mathbf{w^T \Sigma w} \right) \tag{1}$$

$$\sum_{i=1}^{N} w_i = 1 \tag{2}$$

$$0 \leq w_i \leq 1 \quad \forall \; i \tag{3}$$

$$\mathbf{w^T r} = \sum_{i=1}^{N} w_i r_i = R_0 \; , \tag{4}$$

Kuvassa 1 on esitetty kuvitteellisista instrumenteista A, B, C ja D koostuvien portfolioiden joukko. Tehokasta rintamaa kuvaa jatkuva viiva ja katkoviivat kuvaavat esimerkkeinä pareittain instrumenteista (D ja C), (C ja B) ja (B ja A) koostuvia portfolioita.



Kuva 1: Markowitzin malli

Kuvasta 1 nähdään, että yhdistelemällä instrumentteja voidaan pienentää keskihajontaa pienemmäksi kuin yksittäisillä instrumenteilla, kuitenkaan vaikuttamatta tuotto-odotukseen. Tämä on kovarianssirakenteen ominaisuus, jota Markowitzin malli hyödyntää. Tehokkaassa portfoliossa voi olla esimerkiksi kahden sellaisen yhtiön osakkeita, joista toinen tuottaa energiaa öljyllä ja toinen ydinvoimalla. Yhtiöiden osakkeilla on todennäköisesti negatiivista keskinäistä riippuvuutta (toisen kurssi laskee kun toisen nousee), jolloin salkun kokonaishajonta pienenee. Näin vähennetään esimerkiksi öljyn hinnan ja toisaalta uraanin hinnan hajonnasta sjoittajalle koituvaa riskiä.

3.2 Odotettujen tuottojen estimointi

Tässä työssä tuotto-odotusten ennustemallina käytetään faktorimallia:

$$\hat{r}_{it} = \beta + \sum_{k=1}^{4} a_k f_{ikt}, \tag{5}$$

missä $\hat{r}_{it}$ on instrumentin i tuotto-odotus ajanjaksolla $[t, t+T]$, T ennusteperiodin pituus, β affiinitermi, a_k ennustavan tekijän k kontribuutiokerroin ja f_{ikt} on instrumentin tuottoa ennustava tekijä.

Faktorimallin (5) kertoimet määritetään liitteen A pienimmän neliösumman menetelmällä instrumenttien ja faktorimallin tekijöiden historiatiedoista.

Faktorimallin tekijöinä voidaan käyttää instrumenteista riippuen esimerkiksi hinta-tuotto suhdetta tai osinko-voitto suhdetta. Jos instrumentit voidaan kytkeä läheisesti johonkin maahan, voidaan faktoreina käyttää esimerkiksi väestönkasvua, bruttokansantuotteen muutosnopeutta tai työttömyysprosenttia. Käytetyt faktorit pyritään määrittämään siten, että ne kuvaavat parhaiten instrumenttien kehitystä.

3.3 Kovarianssimatriisin estimointi

Olkoon x_{it} instrumentin i arvo hetkellä t. Instrumentin tulevaisuuden odotettu arvo $\mu_{i(t+T)}$ ennusteperiodin lopussa saadaan kaavalla

$$\mu_{i(t+T)} = \hat{r}_{it} x_{it} \tag{6}$$

, missä $\hat{r}_{it}$ on kuten kaavassa (5).

Osakkeen arvon käyttäytymistä ennusteperiodin aikana mallinnetaan siirretyllä Brownin liikkellä, joka on yleisesti käytetty malli kuvaamaan satunnaisesti etenevän prosessin tulevaisuutta.[2] Siirrettyä Brownin liikettä kuvaa differentiaaliyhtälö (7)

$$dx_{i\tau} = dx_i(\tau) = \frac{\mu_{i(t+T)}}{T} d\tau + \tilde{\sigma}_i \epsilon d\tau, \tag{7}$$

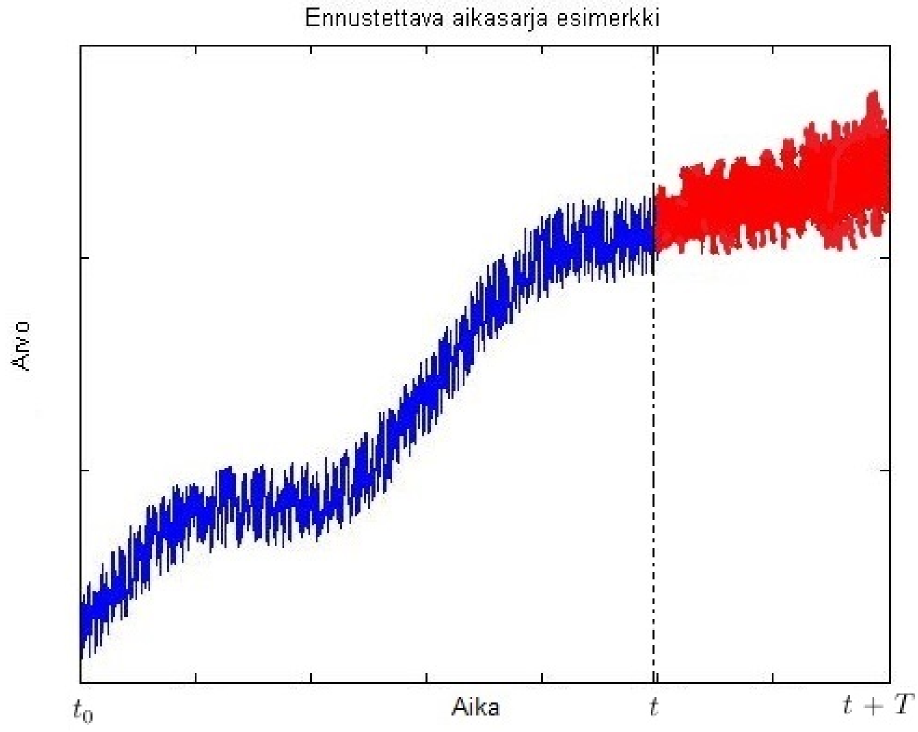
missä $\tilde{\sigma}_i$ on instrumentin i yhden päivän keskihajonta ja $\epsilon \sim N(0,1)$.

Siirretyn Brownin liikkeen yhtälöstä (7) seuraa, että instrumenttien arvot noudattavat hetkellä $\tau > t$ seuraavaa todennäköisyysjakaumaa[2]

$$x_{i\tau} \sim N(\mu_{i\tau}, \Delta t \tilde{\sigma}_i^2) \equiv N\left(x_{it} + \frac{\Delta t}{T}\mu_{i(t+T)}, \Delta t \tilde{\sigma}_i^2\right), \tag{8}$$

missä $\Delta t = \tau - t$.

Kuva 2 hahmottaa siirrettyä Brownin liikettä graafisesti.



Kuva 2: Brownin liike

Nyt kovarianssien estimaatteja ei tarvitse määrittää pelkästään instrumenttien kurssikehityksen historiatiedoista väliltä $[t_0, t]$, vaan niissä voidaan huomioida myös ennusteisiin liittyvä epävarmuus aikavälillä $[t, t+T]$. Kovarianssien estimaatteihin voidaan siis ottaa mukaan kuvan 2 mukainen tulevaisuuden ennuste (ennusteperiodin aikanen kehitys). Brownin liikkeessä on käytetty yhden päivän kovariansseja, koska myös instrumenttien aikasarjoista on otoksia yhdeltä päivältä.

Kovarianssimatriisi estimoidaan siis yhteisesti sekä aikasarjojen tunnetusta historiasta ajalta $[t_0, t]$, että tulevaisuuden Brownin liikkeen avulla. Brownin liike otetaan huomioon jatkuvana ja integroidaan kaikkien tulevaisuuden mahdollisuuksien yli. Saatu eksplisiittinen ratkaisu on tarkempi kuin vaihtoehtoinen menetelmä, tulevaisuuden siirretyn Brownin liikkeen diskreetti Monte-Carlo simulaatio.[5]

Kaava (9) kuvaa matriisimuodossa instrumenttien i ja j kovarianssien estimaatteja $E_{t+T}[\Sigma_{ij}]$ hetkellä $t + T$. Kaavassa (9) r_i on instumentin i yhden päivän tuotto, x_i sen arvovektori ja x_{i0} instrumentin arvo aikasarjan alussa hetkellä t_0. Kovarianssin estimaatissa on mukana integraaliosuus, joka kuvaa siirrettyä Brownin liikettä, todennäköisyystiheysfunktionaan ρ.

$$E_{t+T}[\Sigma_{ij}] = E_{t+T}\left[(r_i - \bar{r}_i)(r_j - \bar{r}_j)\right] = \frac{1}{x_{i0}x_{j0}}\left(E_{t+T}[x_i x_j] - E_{t+T}[x_i]E_{t+T}[x_j]\right) =$$

$$\frac{1}{x_{i0}x_{j0}}\left[\frac{1}{T+t}\left[\underbrace{\sum_{\tau=t_0}^{t} x_{i\tau}x_{j\tau}}_{Historia} + \underbrace{\int_t^{t+T}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i x_j \rho(x_i, x_j, \tau)d\tau dx_i dx_j}_{Brownin\ liike}\right] - E_{t+T}[x_i]E_{t+T}[x_j]\right] \tag{9}$$

Brownin liikkeen oletusten mukaisesti ρ on normaalijakauman tiheysfunktio

$$\rho(x_i, x_j, \tau) = \frac{1}{\sqrt{4\pi^2 det(\tilde{\mathbf{\Sigma}}_{\mathbf{ij}})}} e^{-\frac{1}{2}(\mathbf{x}-\mu_\tau)^{\mathbf{T}}\tilde{\mathbf{\Sigma}}_{\mathbf{ij}}^{-\mathbf{1}}(\mathbf{x}-\mu_\tau)}, \tag{10}$$

missä $\tilde{\mathbf{\Sigma}}_{\mathbf{ij}}$ on instrumenttien i ja j yhden päivän kovariansseista koostuva matriisi.

Seuraavaksi sijoitetaan kaavan (10) lauseke kaavaan (9). Sieventämällä tulos esimerkiksi Mathematican symbolisella laskennalla saadaan

$$E_{t+T}[\Sigma_{ij}] = \frac{1}{x_{i0}x_{j0}}\left[\frac{1}{T+t}\left[\sum_{\tau=t_0}^{t} x_{i\tau}x_{j\tau} + \int_t^{t+T} (\mu_{i\tau}\mu_{j\tau} + \tilde{\sigma}_{ij})d\tau\right] - E_{t+T}[x_i]E_{t+T}[x_j]\right]$$
$$= \frac{1}{x_{i0}x_{j0}}\left[\frac{1}{T+t}\left[\sum_{\tau=t_0}^{t} x_{i\tau}x_{j\tau} + \frac{T}{6}\left[2\mu_i\mu_j + \mu_i x_j + \mu_j x_i + 2x_i x_j\right] + T\tilde{\Sigma}_{ij}\right] - E_{t+T}[x_i]E_{t+T}[x_j]\right]. \tag{11}$$

Instrumenttien arvon oletetaan tulevaisuudessa hetkillä $[t, t+T]$ kasvavan odotusarvoisesti lineaarisesti arvosta x_{it} arvoon $\mu_{i(t+T)}$, josta saadaan instrumenttien odotusarvoiksi

$$E_{t+T}[x_i] = \frac{1}{t+T}\left[\sum_{\tau=t_0}^{t} x_{i\tau} + \frac{T}{2}(\mu_i + x_{it})\right]. \tag{12}$$

Yhdistämällä tulokset (11) ja (12) voidaan laskea kovarianssimatriisin modifioidun estimaatin alkio $E_{t+T}[\Sigma_{ij}]$. Käymällä läpi kaikki indeksiparit (i, j) saadaan laskettua kovarianssimatriisin estimaatti $E_{t+T}[\mathbf{\Sigma}]$.

4 Mallin toimivuuden testaus

4.1 Testiaineisto

Kehitettyjä estimointimenetelmiä testataan valitsemalla tarkasteltaviksi instrumenteiksi maailman likvideimpiä (vaihdetuimpia) USA:ssa noteerattuja ETF-rahastoja (exchange-traded fund). Tarkasteluun valittiin sellaiset rahastot, joiden aikasarjat saatiin täydellisinä ajanjaksolta 1.1.2010-12.6.2015. Sellaiset instrumentit karsittiin pois, joissa oli tapahtunut osakekannan uudelleenjakaminen. Uudelleenjakamisella tarkoitetaan tilannetta, jossa esimerkiksi yksi osake jaetaan kymmeneksi osakkeeksi, joiden arvo on 10 % alkuperäisestä yhden osakkeen arvosta. Tällöin aikasarjasta katsottuna näyttää, että osakkeen arvo tippuu 90 %, vaikka todellisuudessa arvolle ei käy mitään. Tälläisten aikasarjojen poistamisen jälkeen jäljelle jäivät seuraavat 21 ETF-rahastoa:

EZA, EPU, EWA, EWD, EWG, EWI, EWJ, EWM, EWP, EWS, EWT, EWU, EWY, EWQ, EWZ, FXI, IWM, RSX, THD, TUR ja VNM.

Nämä instrumentit jäljittelevät samassa järjestyksessä varsin hyvin seuraavien maiden pörssi-indeksejä:

Etelä-Afrikka, Peru, Australia, Ruotsi, Saksa, Italia, Japani, Malesia, Espanja, Sigmapore, Taiwan, Iso-Britannia, Korea, Ranska, Brasilia, Kiina, USA, Venäjä, Thaimaa, Turkki ja Vietnam.

Koska ETF-rahastot jäljittelevät eri maiden pörssi-indeksejä, pitäisi niiden ajallisen vaihtelun eli keskihajonnan olla huomattavasti pienempää kuin yksittäisillä osakkeilla.

Tässä työssä tuotto-odotukset estimoidaan seuraavaksi esiteltävien tunnuslukujen E/P, B/P, S/P ja Cf/P perusteella, joiden oletetaan ennustavan instrumenttien kurssikehitystä.

$\frac{E}{P}$ tarkoittaa yrityksen tuoton suhdetta pörssiarvoon.

$\frac{B}{P}$ tarkoittaa yrityksen kirja-arvon suhdetta pörssiarvoon.

$\frac{S}{P}$ tarkoittaa yrityksen myynnin suhdetta pörssiarvoon.

$\frac{Cf}{P}$ tarkoittaa yrityksen kassavirran suhdetta pörssiarvoon.

Kyseiset luvut määritetään yleensä vain yksittäisille osakkeille. Kaava (13) havainnollistaa miten tunnusluvut lasketaan tässä työssä useasta osakkeesta koostuvalle rahastolle

$$\frac{X}{P} = \frac{\sum_{n=1}^{N} X_n}{\sum_{n=1}^{N} P_n}, \tag{13}$$

missä $X \in \{E, B, S, Cf\}$ ja P_n on rahastossa olevan yrityksen n osakkeen hinta.

Käytetyt tunnusluvut normeerataan jokaisena ajanhetkenä kunkin 21 instrumentin tunnuslukujen itseisarvojen keskiarvolla kaavan (14) mukaisesti

$$f_{n,normalized} = \frac{f_n}{\sum_{i=1}^{21} |f_i|}, \tag{14}$$

missä $f_n \in \left\{\frac{E}{P}, \frac{B}{P}, \frac{S}{P}, \frac{Cf}{P}\right\}$.

Tämän jälkeen luvut skaalataan vielä seuraavalla funktiolla

$$f_{n,scaled} = a^{f_{n,normalized}}, \text{jossa } a = 1.5. \tag{15}$$

Kaavan (14) normeeraus pitää huolen siitä, että tunnusluvut f skaalautuvat samaan suuruusluokkaan keskenään. Kaavaa (15) on käytetty, koska testaamalla erilaisia regressiomalleja huomataan sen tuottavan paras sovitusaste.

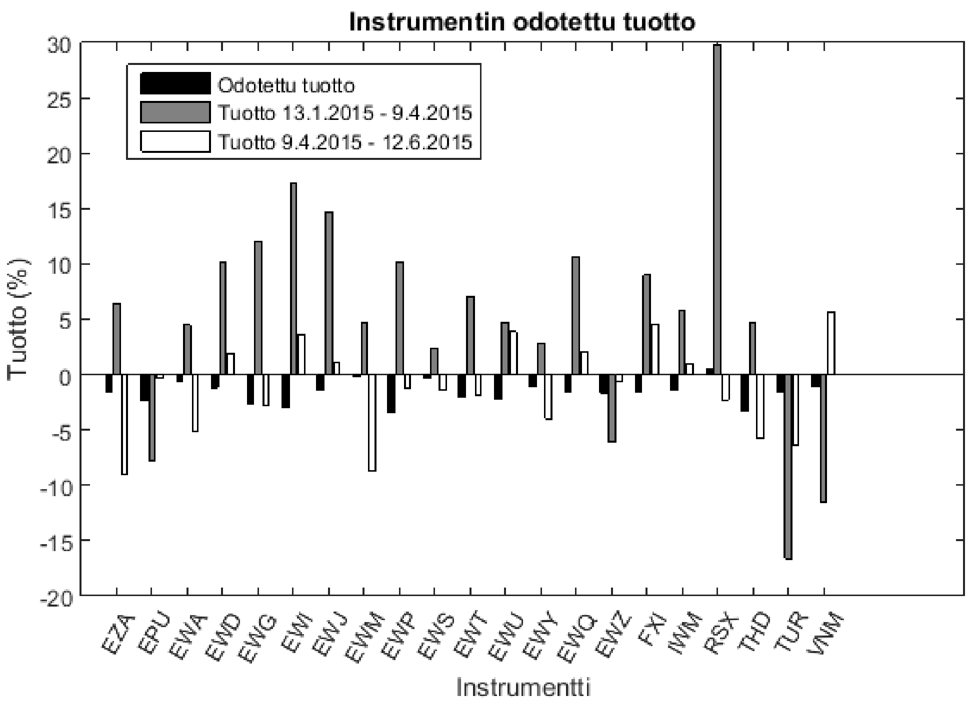
Kun malliin valitut instrumenttien faktorit ja aikasarjat tiedetään ajanjaksolta 1.1.2013 - 12.12.2014, voidaan määrittää kertoimet, joilla skaalatut tunnusluvut kuvaavat parhaiten toteutuneita tuottoja viidellä eri ~ 60 pörssipäivän ajanjaksolla aikana 1.1.2013 - 12.12.2014. Testipisteitä yhtälön (5) faktorimallin parametrien $(a_1, a_2, a_3, a_4, \beta)$ määrittämiseen on siis kokonaisuudessaan $5 \times 4 \times 21$ kpl = 420 kpl. Tuottojen ennuste tehdään tässä työssä ajanjaksoille 13.1.2015 - 9.4.2015 ja 9.4.2015 - 12.6.2015.

Tässä työssä painokertoimet w_i yksittäisille instumenteille on rajattu Markowitzin portfoliomallin yhtälössä (3) enintään arvoon 25%. Tämä takaa sen, että yksittäistä instrumenttia ei tule portfolioihin kohtuuttoman paljon. Tämä on yleinen tapa, jota käytetään portfolion valinnassa pienentämään yksittäisistä instrumenteista koituvaa riskiä.

4.2 Testitulokset

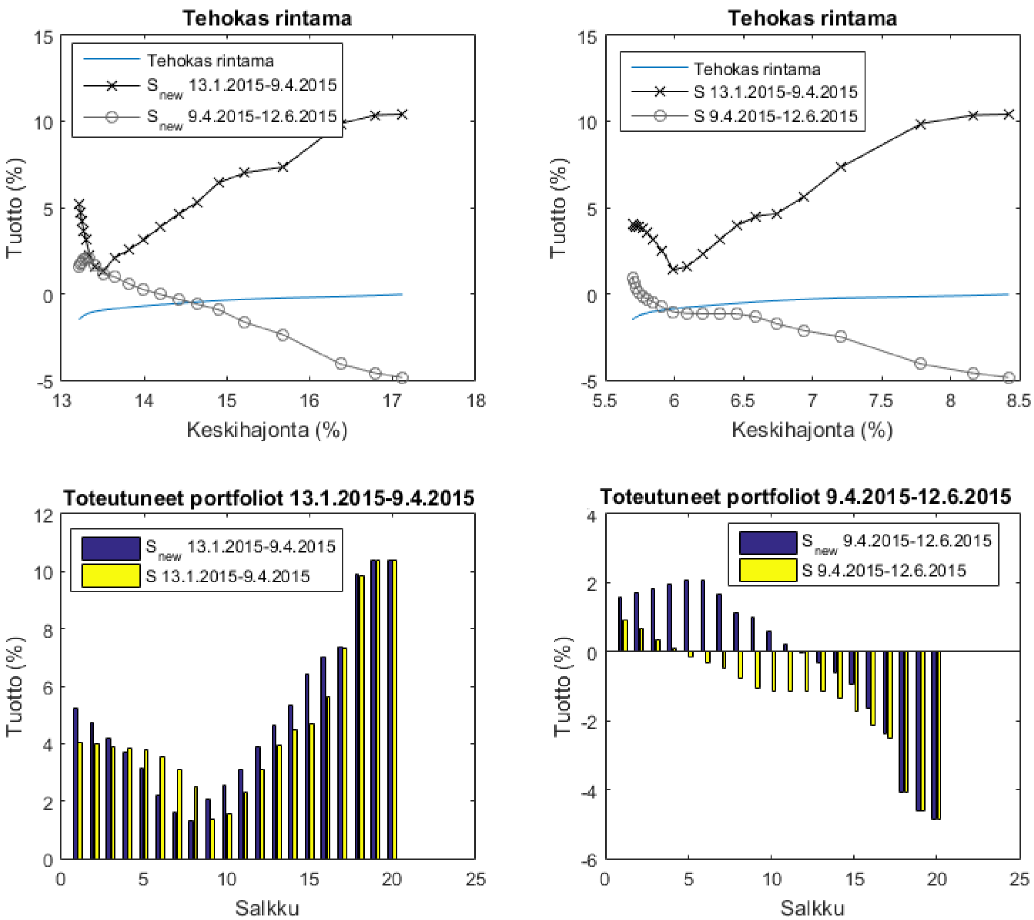
Laskenta on tehty MATLAB R2015b:llä. Tehokkaita portfolioita määritetään eri tuotto-odotuksilla R_0 tasaisesti 20 kpl väliltä $R_0 \in [-1.5\%,\ 0\%]$.

Kuvassa (3) on faktorimallin ennustamat tuotto-odotukset instrumenteille ja instrumenttien toteutuneet tuotot kahdella eri aikaperiodilla.



Kuva 3: Tuotto-odotukset ja toteutuneet tuotot ETF-rahastoille

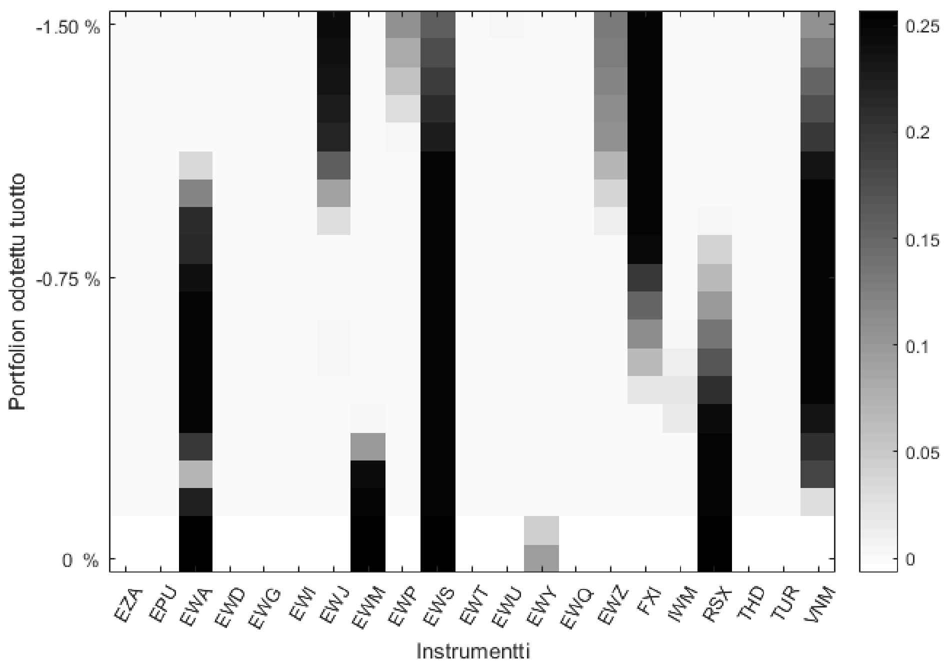
Kuvasta 3 nähdään, että yksittäisten instrumenttien tuotto-odotukset eivät vastanneet toteutuneita tuottoja. Tuotto-odotukset ovat lähellä toteutuneita tuottoja joillekin instrumenteille ja toisilla instrumenteilla ne poikkeavat huomattavasti toisistaan. Faktorimalli tuottojen estimointiin vaikuttaa tällä perusteella heikolta.



Kuva 4: Tehokas rintama tuotto-odotuksilla sekä todellisilla tuotoilla ja portfolioiden (20 kpl) toteutuneet tuotot

Kuvan 4 ylärivin kuvaajissa ovat toteutuneet rintamat modifioidulla kovarianssin estimaatilla (S_{new}) ja vertailun vuoksi klassisella kovarianssilla määritetyt rintamat (S). Brownin liikkeen vuoksi modifioidulla kovarianssin estimaatilla määritettyjen portfolioiden keskihajonnat ovat suurempia. Kuvan 4 alarivin kuvaajissa vertaillaan, miten modifioidulla ja klassisella kovarianssseilla määritetyt tehokkaat portfoliot ovat tuottaneet kahdella eri ajanjaksolla.

Kuvasta 4 nähdään, että tämän työn kovarianssin estimaatti antaa keskimäärin parempia portfolioita kuin perinteinen estimaatti. Perinteinen estimaatti on parempi ainoastaan aikavälin 13.1.2015 - 9.4.2015 salkuilla, joiden toteutuneet tuotot ovat $\sim 4\% - 2\%$, mikä näkyy kuvan 4 kolmannesta kuvaajasta. Ajanjaksolla 9.4.2015 - 12.6.2015 klassinen estimaatti tuottaa aina heikomman portfolion, mikä näkyy kuvan 4 neljännestä kuvaajasta. Suhteellisesti suurilla tuotto-odotuksilla estimaattien antamat portfoliot ovat hyvin samanlaisia. Tämä johtuu siitä, että riippumatta kovariansseista, kuvassa 3 esitetyt matalat tuotto-odotukset pakottavat valitsemaan tiettyjä instrumentteja jos halutaan pitää tuotto-odotus korkeana. Kokonaisuudessaan tämän työn kovarianssin estimaattia voidaan pitää toimivana.



Kuva 5: Painokertoimet eri salkuissa

Kuvassa 5 on estimoitujen tehokkaan rintaman portfolioissa olevien instrumenttien painokertoimet. Painokertoimet eri portfolioiden instrumenteille ovat kuvaajan eri riveillä.

Kuvasta 5 nähdään, että tietyt instrumentit eivät tulleet mukaan yhteenkään portfolioon. Esimeriksi *EZA* (Etelä-Afrikka), *EPU* (Peru) ja *EWD* (Ruotsi) jäävät pois kaikista portfolioista. Instrumenttien pois jääminen johtuu niiden positiivisesta korrelaatiosta muiden instrumenttien kanssa, sekä niille ennustetuista, kuvassa 3 esitetyistä matalista tuotto-odotuksista. Kuvasta 5 nähdään, että esimerkiksi *EWS* (Singapore) on suhteellisesti korkean tuotto-odotuksensa ansiosta mukana kaikissa salkuissa.

5 Yhteenveto ja johtopäätökset

Tässä työssä on kehitetty uusi menetelmä Markowitzin mallin parametrien eli instrumenttien odotettujen tuottojen ja kovarianssien estimointiin. Tuotto-odotuksia ennustettiin faktorimallilla ja instrumenttien keksinäisiä kovariansseja uudentyyppisellä menetelmällä, jossa huomioidaan historiatietojen lisäksi tuleviin tuottoihin liittyvä epävarmuus. Estimaattien toimivuutta testataan soveltamalla kehitettyjä estimointimenetelmiä ETF-rahastoista koostuvien portfolioiden valintaan ja katsomalla, millä tavalla estimaattien perusteella valitut portfoliot kehittyvät tulevaisuudessa.

Tunnuslukujen pohjalta kehitetty tuotto-odotuksia estimoiva faktorimalli vaikuttaa testimateriaalin antamien tulosten perusteella heikolta. Faktorimallin parametrit pitäisi kalibroida suuremmalla määrällä aineistoa ja pidemmällä aikajänteellä, jotta lopullinen toimivuus tulisi testattua luotettavasti. Tämän työn faktorimallin parametrien kalibrointi rajoittui aikaan 2013-2014 ja faktorimallin toimivuuden testaus rajoittui kevääseen 2015. Mallin toteutuneita tuottoja pitäisi katsoa useammalla eri ajankohdalla.

Taloudessa on vuosina 2010-2014 ollut lama-aikaa ja erilaisia heilahteluita, jotka näkyvät ETF-rahastojen kurssikehityksessä. Ennusteena käytettävät tunnusluvut eivät ota näitä asioita huomioon oikealla tavalla. Laman tullessa ETF-rahastojen hinnat laskevat usein ennen kuin vaikutukset ehtivät näkyä yritysten tuloksissa. Tällöin faktorimalli ennustaa hyvää tuottoa vaikka todellisuudessa kurssit alkavat laskea. Tälläinen päätelmä voidaan tehdä tutkimalla kyseisiä ennustavia tunnuslukuja ja kurssikehitystä laman edetessä.

Tässä työssä kovariansseja estimoidaan menetelmällä, joka ottaa huomioon tuotto-odotusten ennustevirheen. Kovarianssin estimoinnissa on mukana Brownin liike, jonka ansiosta pidemmällä ennusteperiodilla portfolioihin tulee mukaan useampia instrumentteja. Tämä voi olla järkevää, sillä pitkällä tähtäimellä se pienentää sijoitusportfolion riskiä. Perinteisellä kovarianssin estimaatilla ilmiötä ei ole, koska estimaatit eivät muutu sijoitusajan pituuden mukaan. Työssä kehitetty menetelmä kovarianssien estimointiin on käyttökelpoinen ainakin noin 1-6 kuukauden ennusteperiodilla.

Tässä työssä kehitetty menetelmä kovarianssin estimointiin osoittautui testiaineistolla toimivaksi ratkaisuksi, joka ottaa huomioon tuotto-odotusten estimointivirheen. Instrumenttien kovarianssin modifioitua estimaattia voidaan hyödyntää myös erilaisilla tuotto-odotusten estimaateilla, kuin mitä tässä työssä käytetään.

6 Kirjallisuusviitteet

[1] *Markowitz H.*, 1592. Portfolio selection, *The Journal of Finance*, Vol. 7, pp. 77-91.

[2] *Hida, T.*, 1980. *Brownian motion*, Springer, United States.

[3] *Polson N., Nicholas G., V. Bernard*, 2000. Bayesian portfolio selection: An empirical analysis of the S&P 500 index 1970-1996, *Journal of Business & Economic Statistics*, Vol. 18., pp. 164-173.

[4] *Berg, B. A.,*, 2008. Introduction to Markov chain Monte Carlo simulations and their statistical analysis, Vol. 9., pp. 11-13.

[5] *Raymond K., G. Zhou*, 2007. Optimal portfolio choice with parameter uncertainty, *The Journal of Financial and Quantitative Analysis*, Vol. 42., pp. 621-656.

[6] *Chan L., J. Karceski, J. Lakonishok*, 1999. On portfolio optimization: Forecasting covariances and choosing the risk model, *The Review of Financial Studies*, Vol.12., pp. 937-974.

[7] *Ledoit O., M. Wolf*, 2003. Improved estimation of the covariance matrix of stock returns with and application of portfolio section, *Journal of Empirical Finance*, Vol. 10., pp. 603-621.

[8] *Garlappi L., T. Wang*, 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach, *The Review of Financial Studies*, Vol. 20., pp. 41-81.

[9] *DeMiguel V., N. Francisco*, 2009. Portfolio Selection with Robust Estimation, *Operations Research*, Vol. 57., pp. 560-577.

[10] *Lopuhaa, P. Hendrik*, 1998. On the relation between S-estimators and M-estimators of multivariate location and covariance, *The Annals of Statistics*, Vol 17., pp. 1662-1683.

7 Liitteet

Liite A

Tämän dokumentin tarkoitusena on määrittää affiinikuvaukseen sellaiset kertoimet, jotka minimoivat virheiden neliöiden summan (pienimmän neliösumman menetelmä).

Affiinikuvauksella tarkoitetaan kuvausta

$\hat{r}_i = a_1 f_{i1} + \ \ldots \ + a_K f_{iK} + \beta = \beta + \sum_{k=1}^{K} a_k f_{ik}$, jossa $\hat{r}_n$ on r_n:n estimaatti a_k ovat regressiomallin kertoimia ja β on vakiotermi.

Tavoitteena on minimoida neliösumma paremetrien $(a_1, ..., a_K, \beta)$ suhteen. Paras regressiomalli neliösummavirheen mielessä saadaan derivoimalla neliövirheiden summaa kaikkien parametrien suhteen

Seuraava kaava esittää minimoitavaa kokonaisvirhettä

$$\sum_{i=1}^{I} \sum_{t=1}^{T} (\hat{r}_{it} - r_{it})^2 = \sum_{i=1}^{I} \sum_{t=1}^{T} [r_{it} - \beta - \sum_{k=1}^{K} a_k f_{ikt}]^2 \Rightarrow$$

$$\min \sum_{i=1}^{I} \sum_{t=1}^{T} [r_{it} - \beta - \sum_{k=1}^{K} a_k f_{ikt}]^2.$$

Derivoidaan a_l suhteen $(1 \leq l \leq K)$ ja etsitään derivaatan nollakohta

$$\frac{\partial}{\partial a_l} \sum_{i=1}^{I} \sum_{t=1}^{T} [r_{it} - \beta - \sum_{k=1}^{K} a_k f_{ikt}]^2 = 0 \quad \Rightarrow$$

$$\sum_{i=1}^{I} \sum_{t=1}^{T} \frac{\partial}{\partial a_l} [r_{it} - \beta - \sum_{k=1}^{K} a_k f_{ikt}]^2 = 0 \quad \Rightarrow$$

$$-2\sum_{i=1}^{I}\sum_{t=1}^{T}[r_{it}-\beta-\sum_{k=1}^{K}a_k f_{ikt}]f_{ilt}=0 \ \ \forall\, l \quad \Rightarrow$$

$$-2\sum_{i=1}^{I}\sum_{t=1}^{T}[f_{ilt}r_{it}-f_{ilt}\beta-f_{ilt}\sum_{k=1}^{K}a_k f_{ikt}]=0 \ \ \forall\, l \quad \Rightarrow$$

$$\sum_{k=1}^{K}a_k\sum_{i=1}^{I}\sum_{t=1}^{T}f_{ikt}f_{ilt}+\beta\sum_{i=1}^{I}\sum_{t=1}^{T}f_{ilt}=\sum_{i=1}^{I}\sum_{t=1}^{T}f_{ilt}r_{it} \ \ \forall\, l \quad \Rightarrow$$

$$\begin{pmatrix} \sum_{t=1}^{T}\sum_{i=1}^{I}f_{i1t}f_{i1t} & \dots & \sum_{t=1}^{T}\sum_{i=1}^{I}f_{i1t}f_{iKt} \\ \dots & \dots & \dots \\ \sum_{i=1}^{I}f_{iKt}f_{i1t} & \dots & \sum_{i=1}^{I}f_{iKt}f_{iKt} \end{pmatrix}\begin{pmatrix} a_1 \\ \dots \\ \dots \\ a_K \end{pmatrix}+\beta\begin{pmatrix} \sum_{i=1}^{I}\sum_{t=1}^{T}f_{i1t} \\ \dots \\ \dots \\ \sum_{i=1}^{I}\sum_{t=1}^{T}f_{iKt} \end{pmatrix}=\begin{pmatrix} \sum_{i=1}^{I}\sum_{t=1}^{T}f_{1kt}r_{it} \\ \dots \\ \dots \\ \sum_{i=1}^{I}\sum_{t=1}^{T}f_{iKt}r_{it} \end{pmatrix}.$$

Saatiin K kappaletta yhtälöitä. Muuttujia ovat a_k, $k=1,...,K$ ja β eli yhteensä $N+1$ muuttujaa. Näiden ratkaisemiseen tarvitaan vielä 1 yhtälö, joka saadaan derivoimalla neliösummaa β:n suhteen

$$\frac{\partial}{\partial\beta}\sum_{i=1}^{I}\sum_{t=1}^{T}[r_{it}-\beta-\sum_{k=1}^{K}a_k f_{ikt}]^2=0 \quad \Rightarrow$$

$$-2\sum_{i=1}^{I}\sum_{t=1}^{T}[r_{it}-\beta-\sum_{k=1}^{K}a_k f_{ikt}]^2=0 \quad \Rightarrow$$

$$\beta=\frac{1}{IT}\sum_{i=1}^{I}\sum_{t=1}^{T}r_{it}-\frac{1}{IT}\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{k=1}^{K}a_k f_{ikt}.$$

Sijoittamalla β alkuperäiseen yhtälöryhmään saadaan lopulta muokkaamalla

$$\begin{pmatrix} a_1 \\ \dots \\ \dots \\ \dots \\ a_K \end{pmatrix}=$$

$$\begin{pmatrix} \left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{i1t}f_{i1t}\right) - \frac{1}{IT}\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{i1t}\right)\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{i1t}\right)\right] & \dots & \left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{i1t}f_{iKt}\right) - ...\right] \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{iKt}f_{i1t}\right) - \frac{1}{IT}\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{iKt}\right)\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{i1t}\right)\right] & \dots & \left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} f_{iKt}f_{iKt}\right) - ...\right] \end{pmatrix}^{-1}$$

$$\times \left(\begin{pmatrix} \sum_{i=1}^{I}\sum_{t=1}^{T} f_{i1t}r_{it} \\ \dots \\ \dots \\ \sum_{i=1}^{I}\sum_{t=1}^{T} f_{iKt}r_{it} \end{pmatrix} - \frac{1}{IT}\left(\sum_{i=1}^{I}\sum_{t=1}^{T} r_{it}\right) \begin{pmatrix} \sum_{i=1}^{I}\sum_{t=1}^{T} f_{i1t} \\ \dots \\ \dots \\ \sum_{i=1}^{I}\sum_{t=1}^{T} f_{iKt} \end{pmatrix} \right).$$

Nämä ovat K-ulotteisen affiinisen regressiomallin kertoimet, jotka minimoivat ennusteiden virheiden neliösumman.