

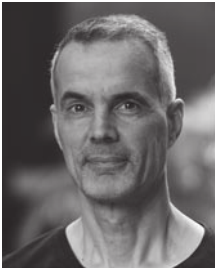


Hoog-dimensionale dynamische copula's voor financiële crises

Tijdens een financiële crisis gaat het merendeel van de financiële instrumenten tegelijk onderuit, waardoor de diversificatie in de markt tijdelijk afneemt. Het modelleren van tijdsafhankelijke, hoog-dimensionale, extreme gebeurtenissen is een complex onderwerp, aangezien elk van deze aspecten afzonderlijk al een uitdaging vormt, laat staan hun combinatie. In dit artikel ontwikkelen we een dynamisch copulamodel, waarbij we recente regularisatietechnieken toepassen om overfitting van de correlatiematrix te voorkomen in hoge dimensies. Als empirische toepassing bestuderen we de Noord-Amerikaanse aandelenmarkt op dagbasis. We stellen vast dat regularisatie en dynamica van de afhankelijkheidsstructuur tot een sterke verbetering leidt van de out-of-sample performance. Met ons dynamische model kunnen we visualiseren hoe de correlatiematrix van de financiële markt zich uitrekt in de gezamenlijke richting tijdens een crisis, wat het systeemrisico verhoogt.

Prof. dr. A. Lucas (links) is hoogleraar financiële econometrie aan de Vrije Universiteit Amsterdam en medebedenker van het Generalized Autoregressive Score model.

Dr. K. Gubbels AAG is actuaaris bij Achmea en academisch directeur van de Executive Master of Actuarial Science van het Actuariel Instituut.



'HONEY, I SHRUNK THE SAMPLE COVARIANCE MATRIX'
Dit is de gevatte titel van een veelgeciteerd artikel door Ledoit en Wolf. Elke actuaaris heeft weleens een correlatiematrix of covariantiematrix geschat op echte data. Als er voldoende datapunten T beschikbaar zijn, en als het aantal variabelen d voldoende klein is (bijvoorbeeld $d/T < 0.1$), dan levert de standaard steekproef-covariantiematrix een goede weergave op van de werkelijke matrix. Echter, als het aantal dimensies toeneemt, dan groeit het aantal parameters in de covariantiematrix kwadratisch, zodat de standaard steekproefschatter onbetrouwbaar wordt. Ledoit en Wolf hebben hun academische levenswerk van dit relevante probleem gemaakt, waarbij ze steeds betere oplossingen hebben ontwikkeld.

Hun meest geavanceerde technieken baseren zich op Random Matrix Theory (RMT). Door middel van RMT is het mogelijk om een asymptotisch exacte relatie vast te stellen tussen het werkelijke eigenwaarden-spectrum en het geschatte spectrum van de steekproef-covariantiematrix. De beroemde wetmatigheid van Marcenko en Pastur stelt bijvoorbeeld dat: als de werkelijke covariantiematrix een eenheidsmatrix is, dan variëren de steekproefeigenwaarden tussen $\lambda_{\pm} = (1 \pm \sqrt{d/T})^2$. Dus stel dat $d/T = 0.5$, dan variëren de steekproef-eigenwaarden van 0.09 tot 2.9, wat zeer grote afwijkingen oplevert van de werkelijke eigenwaarden, die immers gelijk zijn aan 1. Maar we kunnen de relatie ook omdraaien. Stel dat we de geschatte eigenwaarden hebben bepaald: dan kunnen we proberen om de theoretische relatie tussen de steekproefeigenwaarden en de echte eigenwaarden te inverteren om zo het werkelijke onderliggende spectrum terug te vinden. Ledoit en Wolf hebben deze aanpak doorontwikkeld tot hun meest geperfectioneerde variant (aldus Ledoit) die 'quadratic shrinkage' wordt genoemd. Het leidt tot een elegante analytische formule [1], die weliswaar iets te lang is om hier volledig weer te geven, maar die direct toepasbaar is om een robuustere schatting van de covariantiematrix te krijgen in hoge dimensies.

Om de kracht van niet-lineaire shrinkage te illustreren, simuleren we een willekeurige steekproef van 180 observaties (denk aan 15 jaar aan maandelijkse data) voor 100 variabelen (denk aan een intern kapitaalmodel). We nemen voor het gemak aan dat de data multivariaat verdeeld is conform een t -copula met $\nu = 10$ vrijheidsgraden, zodat er sprake is van staartafhankelijkheid. Daarnaast beschouwen we een eenvoudige correlatiestructuur, zoals $\rho_{i,j} = \frac{3}{4} e^{-|i-j|/d}$ voor $i \neq j$. We simuleren twee datasets: 1 om het model (in-sample) te schatten en 1 om de performance van het model (out-of-sample) te meten. In Tabel 1 geven we de in-sample en out-of-sample loglikelihood weer van zowel het echte copulamodel, als het met Maximum Likelihood (ML) geschatte model. We passen ook 'quadratic shrinkage' toe op de geschatte eigenwaarden van de copula-correlatiematrix. In Tabel 1 zien we dat de ML schatter de hoogste in-sample loglikelihood geeft. Dit betreft echter een ernstige vorm van overfitting, wat we kunnen zien aan de out-of-sample performance.

Het toepassen van shrinkage lost de overfitting grotendeels op, wat een enorme out-of-sample performance-winst oplevert.

Loglikelihood	Echt model	ML	Shrinkage
In-sample	10,863	13,926	11,826
Out-of-sample	10,867	6,523	10,396

Tabel 1: In-sample en out-of-sample performance van Maximum Likelihood en Shrinkage schatters voor het gesimuleerde voorbeeld.

GENERALIZED AUTOREGRESSIVE SCORE MODELLEN
Tot dusverre is onze bespreking van de covariantiematrix statisch geweest, terwijl financiële markten in werkelijkheid dynamisch zijn. Een bekend voorbeeld van een dynamisch effect is volatility clustering, wat erop neerkomt dat rustige perioden en volatiele perioden in financiële markten elkaar afwisselen. Het meest gebruikte model om volatility clustering te beschrijven is het GARCH-model, wat de actuaaris bekend voorkomt uit de colleges tijdreeksanalyse. In dit model zorgt een onverwachte schok in het aandelenrendement voor een hogere volatiliteit in navolgende perioden. Iets minder bekende tijdreeksmodellen voor actuarissen zijn de zogenoemde Generalized Autoregressive Score (GAS)-modellen, die een generalisatie vormen van het GARCH-model.

Deze modelfamilie werd geïntroduceerd door Creal, Koopman en Lucas [2], waarbij de afgeleide van de loglikelihood (oftewel, de score) wordt gebruikt om nieuwe informatie uit de markt te verwerken in de parameter-updates. De loglikelihood meet hoe onverwacht de laatste observatie was, gegeven de actuele parameterwaarden. Wanneer de observatie afwijkt van wat het model had verwacht, dan bepaalt de score in welke richting de parameter moet worden aangepast. Door de likelihood-gebaseerde aanpak presteren GAS-modellen vaak beter dan GARCH-modellen. De dynamica van de tijdsafhankelijke modelparameter(s) f_t wordt beschreven door:

$$f_{t+1} = \omega + Af_t + Bs_t.$$

Hierbij geeft s_t de score weer, beschrijft ω het lange-termijn gemiddelde, geeft A de snelheid aan waarmee de parameter terugkeert naar het gemiddelde, en bepaalt B in welke mate de informatie uit de vorige periode wordt meegenomen.

GEREGULARISEERDE DYNAMICA VOOR HOOG-DIMENSIONALE COPULA'S
Nu we hebben vastgesteld hoe we correlatiematrices kunnen verbeteren door deze te regulariseren en dynamisch te maken, passen we deze technieken toe op een hoog-dimensionale copula voor de aandelenmarkt. Het centrale idee achter een dynamische copula is om de totale tijdsafhankelijke verdelingsfunctie te ontbinden in univariate tijdreeksen en marginale verdelingsfuncties voor de individuele variabelen en in de dynamische copula die uitsluitend de samenhang beschrijft.

In onze empirische studie bekijken we dagelijkse logrendementen $r_{i,t}$ voor 100 Noord-Amerikaanse aandelen over een periode van 10 jaar, namelijk van januari 2015 tot 31 december 2024. We passen eerst het standaard GARCH-model toe op elke aandelenreeks en verkrijgen zo de dynamische volatiliteiten $\sigma_{i,t}$. We vervolgen de analyse met de gecorrigeerde rendementen, $\varepsilon_{i,t} = r_{i,t} / \sigma_{i,t}$. Deze hebben univariate verdelingen zonder tijdsafhankelijkheid. De multivariate copula zou echter nog steeds dynamisch kunnen zijn. Om dit te onderzoeken bekijken we de empirische copula-observaties $u_{i,t}$, die we direct kunnen baseren op de rangen van $\varepsilon_{i,t}$. We starten met het schatten

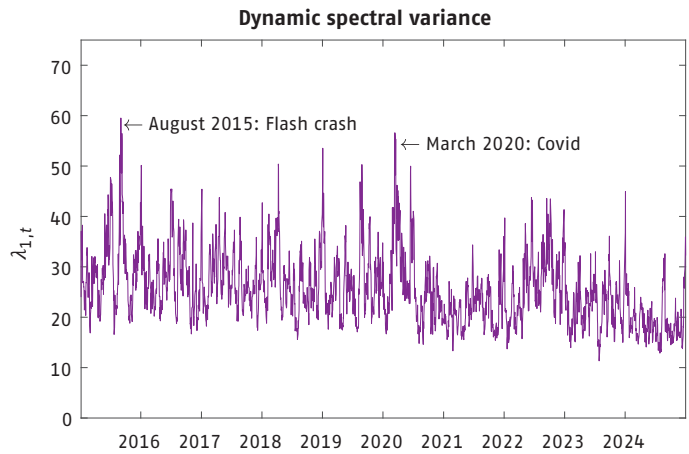
van een standaard t -copula op basis van ML. Dit is een veelgebruikt model voor de samenhang van aandelenrendementen. Tabel 2 toont dat het toepassen van shrinkage op de eigenwaarden van de copula-correlatiematrix de effecten van overfitting reduceert en de out-of-sample performance van het model sterk verbetert.

Loglikelihood	ML	Shrinkage	GAS en shrinkage
In-sample	38,683	38,125	38,604
Out-of-sample	37,376	38,166	38,656

Tabel 2: In-sample en out-of-sample performance van Maximum Likelihood en Shrinkage shatters voor het empirische voorbeeld.

Wanneer we de eigenwaarden van de copula-correlatiematrix vervolgens dynamisch maken met behulp van het GAS-model zien we een verdere verbetering. De eerste dynamische eigenwaarde van deze matrix verklaart de meeste variantie, waarbij de bijbehorende eigenvector de parallelle beweging in de aandelenmarkt weerspiegelt, wanneer alle aandelen tegelijk stijgen of dalen. De resultaten in Tabel 2 laten zien dat het GAS-model met dynamische eigenwaarden tot de beste out-of-sample performance leidt. In Figuur 1 laten we de geschatte dynamica van de eerste eigenwaarde van de correlatiematrix zien. De eigenwaarde ligt onder normale omstandigheden rond de 20, zodat de parallelle richting 20% van de totale variatie in correlatie bepaalt. Echter, in tijden van crisis kan deze waarde toenemen tot wel 60%, wat een navenante toename in systeemrisico en een reductie in diversificatiemogelijkheden tot gevolg heeft.

We eindigen daarom met de conclusie dat zowel regularisatie als dynamica een belangrijke rol spelen bij het modelleren van hoog-dimensionale afhankelijkheden in financiële markten.



Figuur 1: de evolutie van de eerste eigenwaarde van de copula-correlatiematrix over de tijd.

Referenties:
[1] O. Ledoit, M. Wolf (2022), Quadratic shrinkage for large covariance matrices, Bernoulli
[2] D. Creal, S. J. Koopman, A. Lucas (2012), Generalized autoregressive score Models with applications, Journal of Applied Econometrics