C.M.Fortuin,
van Deldenpad 8
6862DC Oosterbeek
fortuinc@xs4all.nl
1998-2005-2007-2015-2017-2019

1 Een bericht in Trouw

Het dagblad Trouw berichtte op 14 augustus 1998 over een onderzoek naar de invloed van de dosis op het effect van chemotherapie bij de nabehandeling van borstkanker. In dat onderzoek kregen 97 borstkankerpatiënten, die voor de chirurgische ingreep allen een chemokuur hadden gekregen, verschillende nabehandelingen: een aantal kreeg een hoge dosis—desondanks kwam bij 19 daarvan de tumor terug; een aantal kreeg een normale dosis—ook hier kwam bij 19 de tumor terug; een aantal van 16 kreeg geen chemotherapie—bij 11 daarvan kwam de tumor terug.

Volgens het bericht was het onderzoeksresultaat, dat een nabehandeling met hoge dosis niet zinvol was. Dat resultaat zou de volgende dag in het tijdschrift 'the Lancet' worden gepubliceerd. Dat tijdschrift lees ik niet (nooit gezien overigens), maar wel was ik nieuwsgierig naar de betekenis van de resultaten. Je geeft nu eenmaal niet voor niets het vak 'statistiek'.

1.1 De correlatie tussen wel en niet chemokuren

Als eerste richt ik me op de vraag hoe effectief, na de chirurgische ingreep, het al of niet behandelen met chemotherapie is (bepaald door het wel of niet terugkeren van de tumor). Daartoe stellen we eerst de aantallentabel op van al of niet 'behandelen' B of $\overline{B}$ tegen al of niet 'genezen' G of $\overline{G}$.

n	G	$\overline{G}$	
B	43	38	81
$\overline{B}$	5	11	16
	48	49	97

Vervolgens benaderen we de kansen P door de relatieve verhoudingen, en berekenen daaruit de varianties en covarianties, bijvoorbeeld

$$\mathbf{var}_B = P_B P_{\bar{B}} \quad \mathbf{covar}_{B,G} = P_{B.G} - P_B P_G$$

Op die manier stellen we de kansentabel en de covariantietabel op

P	G	$\overline{G}$	
B	0,4433	0,3918	0,8351
$\overline{B}$	0,0515	0,1134	0,1649
	0,4948	0,5052	1,0000

covar	G	$\overline{G}$	
B	0,0300	$-0,0300$	0,1377
$\overline{B}$	$-0,0300$	0,0300	0,1377
	0,2500	0,2500	1,0000

Niet toevallig zijn de getallen in de covariantietabel tegengestelden (veroorzaakt door de ontkenning in een van de variabelen). De algemene regel is dat als de vereniging van de gevolgen (oorzaken) alle is, de som van de covarianties in een rij (kolom) 0 moet zijn. In de totaalrij (kolom) staan de kwadraten van de standaard deviaties. Dus berekenen we $\mathbf{covar}_{B,G} = 0,03001$, $S_B = 0,37$ en $S_G = 0,50$. De correlatie tussen behandelen B en genezen G is nu bekend:

$$r_{B,G} = \frac{\mathbf{covar}}{S_B S_G} = \frac{0,03001}{0,37.0,50} = 0,162$$

De berekende correlatie blijkt maar zwak te zijn. Is ze eigenlijk wel significant?

1.2 Correlaties bij normale en hoge dosis

In het krantenbericht stond helaas niet vermeld hoe groot de groepen waren. Om toch verder te kunnen veronderstel ik dat 41 patiënten de normale dosis kregen, en 40 de hoge dosis. Stel de aantallentabel op van 'behandelen normaal' Bn, 'behandelen hoge dosis' Bh of niet behandelen $\overline{B}$ tegen al of niet 'genezen' G of $\overline{G}$ en de bijbehorende kansentabel.

n	G	$\overline{G}$	
Bh	21	19	40
Bn	22	19	41
$\overline{B}$	5	11	16
	28	49	97

P	G	$\overline{G}$	
Bh	0,2165	0,1959	0,4124
Bn	0,2268	0,1959	0,4227
$\overline{B}$	0,0515	0,1134	0,1649
	0,4948	0,5052	1,0000

De covariantietabel en bijpassende correlatietabel worden:

covar	G	$\overline{G}$	
Bh	0,01244	$-0,01244$	0,2423
Bn	0,01765	$-0,01765$	0,2440
$\overline{B}$	$-0,03009$	0,03009	0,1377
	0,2500	0,2500	1,0000

ρ	G	$\overline{G}$
Bh	0,0505	$-0,0505$
Bn	0,0715	$-0,0715$
$\overline{B}$	$-0,1622$	0,1622

Op grond van de correlaties blijkt de enige conclusie die redelijkerwijze te trekken is, zij het ook dan nog voorzichtig, dat niet behandelen niet tot 'genezing' leidt. Of wel behandelen, met welke dosis dan ook, zinvol is, is niet te achterhalen vanwege de bijzonder lage correlaties.

1.3 De significantie van het geen effect hebben

Volgens de standaard theorie kunnen we de significantie toetsen van de alternatieve hypothese: behandeling heeft geen effect, of, genezing is onafhankelijk van behandeling (correlatie nul). Daartoe wordt de χ^2 van de aantallentabel bepaald. In het boek van Andersen en Blankespoor 'Inleiding tot de statistische analyse' wordt deze uitgewerkt tot:

$$\chi^2 = \frac{N(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

waarin $a = n_{BG}$, $b = n_{B\bar{G}}$, $c = n_{\bar{B}G}$, $d = n_{\bar{B}\bar{G}}$. Delen we teller en noemer viermaal door het totaal aantal $N = a+b+c+d$ dan kan dit worden herschreven als:

$$\chi^2 = N\frac{(P_{BG}P_{\bar{B}\bar{G}} - P_{B\bar{G}}P_{\bar{B}G})^2}{(P_{BG}+P_{B\bar{G}})(P_{\bar{B}G}+P_{\bar{B}\bar{G}})(P_{BG}+P_{\bar{B}G})(P_{B\bar{G}}+P_{\bar{B}\bar{G}})}$$

De noemer is gemakkelijk te vereenvoudigen, omdat $P_{BG} + P_{B\bar{G}} = P_B$, enzovoort, tot de vorm:

$$\text{noemer} = (P_B P_{\bar{B}})(P_G P_{\bar{G}}) = S_B^2 S_G^2$$

De factor in de teller is niets anders dan de covariantie tussen B en G:

$$\begin{aligned}
\mathbf{covar}_{B,G} &= P_{BG} - P_B P_G \\
&= P_{BG} - (P_{BG}+P_{B\bar{G}})(P_{BG}+P_{\bar{B}G}) \\
&= P_{BG}(1 - P_{BG} - P_{B\bar{G}} - P_{\bar{B}G}) - P_{B\bar{G}}P_{\bar{B}G} \\
&= P_{BG}P_{\bar{B}\bar{G}} - P_{B\bar{G}}P_{\bar{B}G} \\
\text{teller} &= \mathbf{covar}^2_{B,G} = r^2_{B,G} S_B^2 S_G^2
\end{aligned}$$

Op grond hiervan kan de significantiefunctie χ^2 rechtstreeks in het aantal waarnemingen en de correlatie worden uitgedrukt [1]:

$$\boxed{\chi^2 = N r^2_{B,G}}$$

[1] Dit was al bekend aan R. A. Fisher, zoals te lezen is in zijn artikel over 'on the interpretation of χ^2 from contingency tables, and the calculation of P'(1922); zie naschrift.

Met $N = 97$ en $r = 0,162$ is $\chi^2 = 2,55$. Het aantal vrijheidsgraden van de tabel is $\nu = 1$. In dat geval is de t-verdeling voor χ^2 gelijk aan de normale verdeling voor χ. Bij $\chi^2 = 2,55$ hoort $\chi = u = 1,596$, met een eenzijdige overschrijdingskans van 0,055. De correlatie kan tweezijdig afwijken van 0; er is dus 2.5,5=11% kans dat er geen correlatie is tussen behandelen en genezen. Omgekeerd is er dus 89% kans dat behandelen wel degelijk tumoren onderdrukt.

1.4 De significantie van het wel effect hebben

We gaan nog even op dezelfde voet door, maar nu toetsen we de hypothese dat de behandeling een correlatie ρ heeft. Helaas laat het boek van Anderson en Blankespoor ons nu in de steek, zodat we de formule zelf moeten uitvogelen. Voor het bepalen van χ^2 wordt het 'gevonden' aantal vergeleken met het 'theoretische' aantal. De verschillende mogelijke gevallen worden dan op een 'bepaalde manier' gesommeerd:

$$\chi^2 = (n_{BG} - n_{BG}^{\text{th}})^2 / n_{BG}^{\text{th}} + \cdots$$

Het (te verwachten) aantal is evenredig met de (onbekende) theoretische kans P^{th}:

$$\chi^2 = N(P_{BG} - P_{BG}^{\text{th}})^2 / P_{BG}^{\text{th}} + \cdots$$

Volgens onze hypothese geldt:

$$\mathbf{covar}_{BG}^{\text{th}} = \rho_{B,G} S_B S_G$$

waarmee de kans op 'behandelen met succes' theoretisch wordt:

$$P_{BG}^{\text{th}} = P_B P_G + \rho_{B,G} S_B S_G$$

Het verschil tussen 'praktijk' en 'theorie' (voor dit geval) wordt daarmee:

$$\begin{aligned}
(P_{BG} - P_{BG}^{\text{th}}) &= P_{BG} - P_B P_G - \rho_{B,G} S_B S_G \\
&= \mathbf{covar_{B,G}} - \rho_{B,G} S_B S_G \\
&= r_{B,G} S_B S_G - \rho_{B,G} S_B S_G \\
&= (r_{B,G} - \rho_{B,G}) S_B S_G
\end{aligned}$$

Het verschil tussen de *aantallen (kansen)* van theorie en praktijk is op een factor na gelijk aan het verschil van de *correlaties*; ter herinnering: de correlatie r wordt praktisch 'gemeten', de waarde ρ theoretisch 'verwacht'. Passen we dit resultaat toe op de andere gevallen (met ontkenningen), dan geeft iedere ontkenning eenvoudig 'een extra min'; bijvoorbeeld

$$(P_{\bar{B}G} - P_{\bar{B}G}^{\text{th}}) = (r_{\bar{B},G} - \rho_{\bar{B},G}) S_{\bar{B}} S_G = -(r_{B,G} - \rho_{B,G}) S_B S_G$$

Invullen in chi-kwadraat van alle gevallen (kort $r_{B,G}$ af tot r, $\rho_{B,G}$ tot ρ) geeft:

$$\chi^2 = N(r-\rho)^2 S_B^2 S_G^2 \left(\frac{1}{P_{BG}^{\text{th}}} + \frac{1}{P_{\bar{B}G}^{\text{th}}} + \frac{1}{P_{B\bar{G}}^{\text{th}}} + \frac{1}{P_{\bar{B}\bar{G}}^{\text{th}}} \right)$$

Verdere vereenvoudiging ontstaat door de som der omgekeerden onder een noemer te brengen en die uit te werken naar machten van ρ. Het resultaat daarvan kan hier achter worden nagelezen.[2] Gecombineerd met wat we al hadden wordt chi-kwadraat:

$$\begin{aligned}
\chi^2 = \; & N\left(\frac{r-\rho}{1-\rho^2}\right)^2 \Big/ \left[1 + \frac{\rho}{(1+\rho)^2} \frac{(P_B - P_{\bar{B}})}{S_B} \frac{(P_G - P_{\bar{G}})}{S_G} - \right. \\
& \left. -\left(\frac{\rho}{1-\rho^2}\right)^2 \left(\frac{(P_B - P_{\bar{B}})}{S_B} - \frac{(P_G - P_{\bar{G}})}{S_G} \right)^2 \right]
\end{aligned}$$

Merk op, dat voor $\rho = 0$ (gelukkig) de chi-kwadraat van de onafhankelijkheidstest terugkomt. Verwaarlozen van de verschiltermen ($P_G \approx 1/2$) geeft een benaderende uitdrukking voor chi-kwadraat voor $\rho \approx 0$:

$$\boxed{\chi^2 \approx N((r-\rho)/(1-\rho^2))^2}$$

[2] Zie Appendix: A. Uitwerking: de som der omgekeerden.

Dit betekent dat de correlatie rond de gevonden waarde r een betrouwbaarheidsgebied heeft bepaald door $\pm(1-\rho^2)\chi/\sqrt{N}$. Met $\chi = 1,96$ heeft het 95% betrouwbaarheidsgebied rond $r = 0,162$ de breedte $\pm 0,194$; rekening houdend met $\frac{P_B - P_{\bar{B}}}{S_B} = 65/36$ wordt dat $\pm 0,189$ (daar ligt dus ook 0 in). De waarde 0 van de correlatie kunnen we uitsluiten door de iets lagere betrouwbaarheid van 89% te verlangen.

2 Significantie bij onafhankelijkheid maar meerdere oorzaken en gevolgen

Kunnen we deze analyse naar meerdere gevallen uitbreiden? In het kranteberricht was sprake van drie soorten doses—namelijk hoge, normale en geen dosis van de chemokuur—en van twee soorten effect: wel of niet terugkeren van de tumor. We zullen gemakshalve in het algemene geval de verschillende oorzaken aanduiden met i genummerd $1 \cdots k$, de verschillende gevolgen aanduiden met j, genummerd $1 \cdots l$. Met $\sum'$ wordt de sommatie over alle behalve de laatste bedoeld: ${\sum'}_i = \sum_{i=1}^{k-1}$, ${\sum'}_j = \sum_{j=1}^{l-1}$.

Omdat de verschillende i, resp j, samen alle mogelijkheden bevatten gelden beperkingen op de kolommen, resp rijen, van de tabellen. Voor de kanstabel geldt dat de som van alle rijen en kolommen 1 is. Voor de covariantietabel geldt dat de som van een rij, of kolom, 0 is. Voor de correlatietabel geldt dat de gewogen som van een rij, met gewicht S_j, of een kolom, met gewicht S_i, nul is. Op grond van deze somregels kunnen we herschrijven:

$$\begin{array}{lll} P_k = 1 - \sum_i{}' P_i & P_{kj} = P_j - \sum_i{}' P_{ij} & \mathbf{covar}_{kj} = -\sum_i{}' \mathbf{covar}_{ij} \\ P_l = 1 - \sum_j{}' P_j & P_{il} = P_i - \sum_j{}' P_{ij} & \mathbf{covar}_{il} = -\sum_j{}' \mathbf{covar}_{ij} \end{array}$$

De *hypothese* die we zullen testen is dat oorzaken en gevolgen *onderling onafhankelijk* zijn. Dat houdt in voor de kansen, dat:

$$\begin{aligned} P_{ij}^{\text{th}} &= P_i P_j \\ (P_{ij} - P_{ij}^{\text{th}}) &= \mathbf{covar}_{ij} \end{aligned}$$

De term in chi-kwadraat die overeenkomt met het geval ij is:

$$\chi_{ij}^2 = \frac{(n_{ij} - n_{ij}^{\text{th}})^2}{n_{ij}^{\text{th}}} = N\frac{(P_{ij} - P_{ij}^{\text{th}})^2}{P_{ij}^{\text{th}}} = N\frac{\mathbf{covar}_{ij}^2}{P_i P_j}$$

De sommatie over alle gevallen, die nodig is voor het bepalen van chi-kwadraat, voeren we in stappen uit: eerst de sommatie over i, de rijen, dan de sommatie over j, de kolommen. De sommatie over de rijen van chi-kwadraat termen levert (we laten bij de covariantie 'praktijk' weg):

$$\begin{aligned} \chi_j^2 = N \sum_i & \frac{(\mathbf{covar}_{ij})^2}{P_i P_j} \\ &= \frac{N}{P_j}\left(\sum_i{}' \frac{(\mathbf{covar}_{ij})^2}{P_i} + \frac{(\mathbf{covar}_{kj})^2}{P_k}\right) \\ &= \frac{N}{P_j}\left(\sum_i{}' \frac{(\mathbf{covar}_{ij})^2}{P_i} + \frac{(-\sum_i{}' \mathbf{covar}_{ij})^2}{P_k}\right) \\ &= \frac{N}{P_j}\left(\sum_i{}' \sum_{i'}{}' \delta_{i,i'} \frac{\mathbf{covar}_{ij}\mathbf{covar}_{i'j}}{P_i} + \frac{\sum_i{}' \sum_{i'}{}' \mathbf{covar}_{ij}\mathbf{covar}_{i'j}}{P_k}\right) \\ &= \frac{N}{P_j} \sum_i{}' \sum_{i'}{}' \mathbf{covar}_{ij}\left(\delta_{i,i'}\frac{1}{P_i} + \frac{1}{P_k}\right)\mathbf{covar}_{i'j} \end{aligned}$$

waarbij gebruik is gemaakt van de gelijkheidsindicator $\delta_{i,i'}$ om de sommatie van alle termen overeen te doen stemmen.
We hebben te maken met een *symmetrische tweedegraads vorm* met bijbehorende *symmetrische matrix*

$$\left(F = \delta_{i,i'}\frac{1}{P_i} + \frac{1}{P_k}\right)$$

In dit geval wordt de matrix F *alleen* door de oorzaken i bepaald. Volgens de standaard theorie kan een symmetrische matrix op diagonaalvorm worden gebracht met behulp van de eigenwaarden en genormeerde eigenvectoren die passen bij de symmetrische matrix. Daarbij wordt overgegaan van een beschrijving van de matrix door de oorzaken i naar de gestandaardizeerde beschrijving van de matrix door oorzaken λ, die lineaire combinaties zijn van de oorzaken i. Het bijzondere is, dat deze 'oorzaken λ' statistisch *onafhankelijk* zijn, in tegenstelling tot de oorzaken i, die statistisch wel afhankelijk zijn (juist omdat ze elkaar uitsluiten!): de gestandaardizeerde oorzaken λ zijn 'onafhankelijke oorzaken'.

De getallen $1/\lambda$ (de omkering is om dimensionele redenen zo gekozen) zijn de omgekeerde eigenwaarden van de symmetrische matrix. De bijbehorende eigenvectoren zijn $v_{i\lambda}$. De $k-1$ getallen $v_{i\lambda}$ zijn de gewichten (mogelijk negatief) van de oorzaken i waaruit de 'onafhankelijke oorzaak λ' is samengesteld. We zullen apart zien dat de $k-1$ getallen λ, op een factor na, de varianties zijn op de bijpassende 'onafhankelijke oorzaken' λ.

2.1 De onafhankelijke oorzaken λ

Eerst laten we zien hoe de gewichten $v_{i\lambda}$ kunnen worden bepaald, en hoe de λ kunnen worden gevonden. Krachtens de definitie van eigenwaarde en eigenvector moet de eigenvectorvergelijking gelden:

$$\sum_{i'}{}' \left(\delta_{i,i'}\frac{1}{P_i} + \frac{1}{P_k}\right) v_{i'\lambda} = \left(\frac{v_{i\lambda}}{P_i} + \frac{\sum'_{i'} v_{i'\lambda}}{P_k}\right) = v_{i\lambda}/\lambda$$

We definiëren als extra variabele de som over de eigenvectorcomponenten:

$$v_\lambda = \sum_i{}' v_{i\lambda}$$

De oplossing voor de *eigenvector* met gebruikmaking van de somvariabele is:

$$\boxed{v_{i\lambda} = \frac{1}{1/\lambda - 1/P_i}\frac{v_\lambda}{P_k}}$$

Substitutie van $v_{i\lambda}$ in somvariabele v_λ, nodig voor consistentie, leidt, ongeacht de waarde van v_λ, tot de *eigenwaardenvergelijking* voor de λ's:

$$\sum_i{}' \frac{1}{1/\lambda - 1/P_i} = P_k = 1 - \sum_i{}' P_i \Rightarrow \boxed{\sum_i{}' \frac{P_i^2}{P_i - \lambda} = 1}$$

De eigenvector wordt vastgelegd door de waarde van v_λ. Gebruikelijk is lengte 1: $\sum_i{}' v_{i\lambda}^2 = 1$, of $1/v_\lambda^2 = \sum_i{}' 1/P_k^2/(1/\lambda - 1/P_i)^2 = (\lambda/P_k)^2 \sum_i{}' P_i^2/(P_i - \lambda)^2$.

$$\boxed{v_\lambda = \frac{P_k}{\lambda} / \sqrt{\sum_i{}' P_i^2/(P_i - \lambda)^2}}$$

Het nul worden van $y = \sum_i{}' P_i^2/(P_i - \lambda) - 1$ bepaalt de eigenwaarde. [3] De functie y van λ valt van nature uiteen in open intervallen, door het onbepaald zijn als de noemer nul is: $(-\infty, P_1)$, (P_1, P_2), (P_2, P_3), $\ldots (P_{k-2}, P_{k-1})$, (P_{k-1}, ∞), gesteld dat de kansen P_i geordend zijn naar stijgende grootte. In ieder interval is y stijgend. Bij een linker P_i-randpunt wordt de functie

[3] dit is feitelijk gelijk aan het 'nul worden van de determinant' van de eigenwaardematrix

$-\infty$, Bij een rechter P_i-randpunt wordt de functie $+\infty$. In ieder open (P_{i-1}, P_i)-interval is er dus precies één eigenwaarde λ_i waar y nul is. Bij het meest linkse interval is bij $\lambda = 0$ de $y = \sum_i{}' P_i - 1 = -P_k$, dus negatief. Voor de kleinste eigenwaarde λ_1 geldt dus $0 < \lambda_1 < P_1$. Op het meest rechtse interval eindigt y met -1, negatief: daar is geen eigenwaarde. Totaal zijn er $k-1$ eigenwaarden. Samenvattend:

$$0 < \lambda_1 < P_1 < \lambda_2 \cdots < P_{i-1} < \lambda_i < P_i < 1$$

Merk op, dat voor ieder samenvallende P_i een eigenwaarde verdwijnt. Verder lijken de eigenwaarden af te hangen van de keuze van 'k'. De som over àlle i: $\Sigma_i P_i^2/(P_i-\lambda) = 1 + P_k^2/(P_k-\lambda)$ is P_k-afhankelijk (?).

Voorbeeld 2 oorzaken ($k = 2$): Er is slechts één onafhankelijke oorzaak 1 en één eigenwaarde $\lambda_1 = P_1 - P_1^2 = S_1^2 = P_1 P_k < P_1$, en één 'eigenvector' van één getal $v_{1\lambda} = (1/P_k)/(1/(P_1 P_k) - 1/P_1) v_\lambda = v_\lambda [= 1]$.

Voorbeeld 3 oorzaken ($k = 3$): Er zijn 2 onafhankelijke oorzaken en twee eigenwaarden λ. De eigenwaardevergelijking wordt een tweedegraadsveelterm:

$$\begin{aligned}
1 &= P_1^2/(\lambda - P_1) + P_2^2/(\lambda - P_2) \\
(\lambda - P_1)(\lambda - P_2) &= P_1^2(\lambda - P_2) + P_2^2(\lambda - P_1) \\
0 &= \lambda^2 - (P_1 + P_2 - P_1^2 - P_2^2)\lambda + P_1 P_2 (1 - P_1 - P_2) \\
0 &= \lambda^2 - (P_1(1-P_1) + P_2(1-P_2))\lambda + P_1 P_2(1 - P_1 - P_2) \\
\mathbf{discr} &= (P_1(1-P_1) + P_2(1-P_2))^2 - 4P_1P_2(1-P_1-P_2) \\
&= (P_1(1-P_1) - P_2(1-P_2))^2 + 4P_1P_2((1-P_1)(1-P_2) - (1-P_1-P_2)) \\
&= (P_1 - P_2)^2(1 - (P_1 + P_2))^2 + 4P_1^2P_2^2 \geq (2P_1P_2)^2
\end{aligned}$$

Bij ons oorspronkelijke bestralingsprobleem zijn $P_1 = 0,4124$ en $P_2 = 0,4227$. De eigenwaarden voldoen aan $0 < \lambda_1 = 0,0689 < P_1 < \lambda_2 = 0.4175 < P_2$. Merk op, dat $\lambda_2 - \lambda_1 = 0,3486 = \sqrt{\mathbf{discr}} \geq 2P_1P_2 = 0,3486.$(?!)

Tweedens bepalen we de kans op 'onafhankelijkheid' van de oorzaken 'λ'. Omdat ze oorzaak k niet bevatten is $v_{k\lambda} = 0$, zodat alleen $i \neq k$ telt. De kans P_λ op 'oorzaak λ' is:

$$\begin{aligned}
P_\lambda &= \sum_i P_i v_{i\lambda} = \sum_i P_i \frac{1}{1/\lambda - 1/P_i} \frac{v_\lambda}{P_k} = \lambda \frac{v_\lambda}{P_k} \sum_i \frac{P_i^2}{P_i - \lambda} \\
P_\lambda &= \lambda \frac{v_\lambda}{P_k}
\end{aligned}$$

Bijvoorbeeld 2 oorzaken ($k = 2$): $v_\lambda = 1$ en de kans op 'de' onafhankelijke oorzaak is $P_\lambda = P_1 P_k v_\lambda / P_k = P_1 v_\lambda = P_1$. Merk op dat de eigenvector kan worden herschreven zonder v_λ maar met P_λ:

$$v_{i\lambda} = \frac{P_\lambda/\lambda}{(1/\lambda - 1/P_i)} = \frac{P_i P_\lambda}{P_i - \lambda}$$

De onderlinge onafhankelijkheid van de oorzaken λ en λ' blijkt uit het nul zijn van de covarianties als $\lambda \neq \lambda'$, vanwege de bijzondere orthonormale eigenschap van de eigenvectoren van symmetrische matrices:

$$\begin{aligned}
\mathbf{covar}_{\lambda,\lambda'} &= \sum_i{}' P_i v_{i\lambda} v_{i\lambda'} - P_\lambda P_{\lambda'} \\
&= \sum_i{}' \lambda(1 + P_i(1/\lambda - 1/P_i)) v_{i\lambda} v_{i\lambda'} - P_\lambda P_{\lambda'} \\
&= \lambda \sum_i{}' (v_{i\lambda} + P_i P_\lambda/\lambda) v_{i\lambda'} - P_\lambda P_{\lambda'} \\
&= \lambda \sum_i{}' v_{i\lambda} v_{i\lambda'} + P_\lambda P_{\lambda'} - P_\lambda P_{\lambda'} = \lambda \delta_{\lambda,\lambda'}
\end{aligned}$$

Dat $v_{i\lambda'}$ en $v_{i\lambda}$ orthogonaal zijn volgt rechtstreeks uit het voldoen van $\lambda' \neq \lambda$ als eigenwaarden aan de eigenwaardevergelijking;

$$v_{i\lambda}v_{i\lambda'} = \frac{P_i P_\lambda}{P_i - \lambda}\frac{P_i P_{\lambda'}}{P_i - \lambda'} = P_\lambda P_{\lambda'}\frac{1}{\lambda - \lambda'}\left(\frac{P_i^2}{P_i - \lambda} - \frac{P_i^2}{P_i - \lambda'}\right) \rightarrow \sum_i{}' v_{i\lambda}v_{i\lambda'} = 0$$

Als $\lambda' = \lambda$ dan wordt de covariantie de variantie S_λ^2 van 'oorzaak λ':

$$S_\lambda^2 = \lambda$$

Bij slechts 2 oorzaken ($k = 2$), is de variantie $S_\lambda^2 = \lambda^2 = P_1 P_k = P_1 P_2 = P_1(1-P_1)$ zodat inderdaad $S_\lambda = S_1$.

2.2 Chi-kwadraat bij onafhankelijkheid van oorzaken en gevolgen

Nu we hebben achterhaald hoe om te gaan met de 'onafhankelijke oorzaken' kunnen we de bijdrage tot chi-kwadraat bepalen. De covariantie van de 'oorzaak λ' met gevolg j wordt:

$$\mathbf{covar}_{\lambda j} = \sum_i{}' v_{i\lambda}\mathbf{covar}_{ij}$$

De chi-kwadraat bijdrage van gevolg j wordt na genormeerde v_λ diagonalistie:

$$\chi_j^2 = \frac{N}{P_j}\sum_\lambda{}' \frac{\mathbf{covar}_{\lambda j}^2}{\lambda} = N\sum_\lambda{}' \frac{\mathbf{covar}_{\lambda j}^2}{S_\lambda^2 P_j}$$

Merk op, dat de vorm van de termen in de som niet is gewijzigd door de overgang naar de onafhankelijke oorzaken, behalve dan dat de kans is vervangen door de variantie!
Precies zoals we de oorzaken (rijen i) hebben behandeld, kunnen we ook de gevolgen (kolommen j) behandelen. Daarbij is het van belang op te merken, dat de kolomregel ook geldt voor de covarianties $\mathbf{covar}_{\lambda j}$ van de 'oorzaken λ' (omdat die een lineaire combinatie van gewone oorzaken zijn). We kunnen netzo tot 'onafhankelijke gevolgen μ' komen met als gewichten de eigenvectoren $v_{\mu,j}$ en de eigenwaardenvergelijking:

$$v_{\mu,j} = \frac{P_j P_\mu}{P_j - \mu} \qquad 0 = \sum_j{}' \frac{P_j^2}{P_j - \mu} - 1$$

De chi-kwadraatsom wordt daarmee:

$$\begin{aligned}\chi^2 &= \sum_j{}' \chi_j^2 = \sum_j{}' N\sum_\lambda{}' \frac{\mathbf{covar}_{\lambda j}^2}{S_\lambda^2 P_j} = N\sum_\lambda{}' \frac{1}{S_\lambda^2}\sum_j{}' \frac{\mathbf{covar}_{\lambda j}^2}{P_j} \\ &= N\sum_\lambda{}'\sum_\mu{}' \frac{\mathbf{covar}_{\lambda,\mu}^2}{S_\lambda^2 S_\mu^2}\end{aligned}$$

Na veel omwerking hebben we een treffend eenvoudige formule:

$$\boxed{\chi^2 = N\sum_{\lambda,\mu}{}'' r_{\lambda,\mu}^2}$$

Bijvoorbeeld, bij 2 oorzaken en 2 gevolgen ($k = 2$ en $l = 2$) is er slechts één onafhankelijke correlatie en wordt $\chi^2 = Nr^2$. Iedere 'onafhankelijke oorzaak'-'onafhankelijke gevolg' combinatie draagt een term bij gelijk aan het kwadraat van de onderlinge correlatie. Daarom is in het bijzondere geval dat oorzaken en gevolgen dezelfde zijn het resultaat eenvoudig: $\chi^2 = N(k-1)$.

[2017] In navolging van Hotelling (1936)[4] kunnen we een stap verder gaan. Een algemene 'matrix' kan worden 'gediagonaliseerd'; hier $A = r_{\lambda,\mu}^2$. De diagonaaltermen, zijn hier eveneens

[4] zie Naschrift: Canonieke variabelen volgens Hotelling

getallen tussen 0 en 1, en dus als verschillende onafhankelijke r^2 te beschouwen. De *algemene theorie* is als volgt: Vat A_{ij} op als matrix A tussen de ruimten i links en j rechts. Dan is er een (gedeeltelijk) diagonale $\hat{A}$ met eigenwaarden α op de diagonaal van de ij ruimte. Op de diagonaal zien we de koppeling van de i en j eigenvectoren die bij elkaar horen— losjes gezegd is voor die vectoren $i = j$ (*canonieke variabelen*). De linker eigenvectoren vormen een orthonormale rotatie L^t in de i ruimte, de rechter eigenvectoren vormen een orthonormale rotatie R in de j ruimte:

$$L^t A = \hat{A} R^t \leftrightarrow A = L\hat{A}R^t \leftrightarrow AR = L\hat{A},$$

waarbij de (verschillende) α's eigenwaarden zijn van de linker eigenvectoren als kolommen van L, of van de rechter eigenvectoren als kolommen van R. Vanwege dat eigenvector karakter is $AA^tL = L\hat{A}\hat{A}^t = L\alpha\alpha^*$ en $A^tAR = R\hat{A}^t\hat{A} = R\alpha^*\alpha$, hetwelk de bepaling van de eigenwaarden α^2 en en -vectoren L en R terugbrengt tot die van gewone vierkante symmetrische matrices.

3 Significantie bij afhankelijkheid van meerdere oorzaken en gevolgen

Aangemoedigd door het succes van het gebruik van de 'onafhankelijke oorzaken en gevolgen' richten we ons nu op de situatie dat oorzaak en gevolg afhankelijk zijn. De *hypothese* die we testen is gebaseerd op onderlinge *afhankelijkheid van oorzaak en gevolg* volgens:

$$P^{\mathrm{h}}_{ij} = P^{\mathrm{h}}_i P^{\mathrm{h}}_j + \mathbf{covar}^{\mathrm{h}}_{ij},$$

De kansen P^{h} en covarianties $\mathbf{covar}^{\mathrm{h}}_{ij}$ worden bepaald door het $h$ypothesemodel: de hypothetische kansen en varianties van de oorzaak en gevolg variabelen zijn gelijk aan de praktische kansen en variaties. Wiskundig geformuleerd is het hypothesemodel:

$$P^{\mathrm{h}}_i = P_i, \quad S^h_i = S_i, \quad P^{\mathrm{h}}_j = P_j, \quad S^h_j = S_j.$$

Daaruit volgt voor de *excesvariabelen* e, de afwijkingen van de kans van die van de hypothese:

$$e_{ij} = (P_{ij} - P^{\mathrm{h}}_{ij}) = \mathbf{covar}_{ij} - \mathbf{covar}^{\mathrm{h}}_{ij}$$

De term in chi-kwadraat die overeenkomt met het element ij is:

$$\chi^2_{ij} = N\frac{e^2_{ij}}{P^{\mathrm{h}}_{ij}} = N\frac{(\mathbf{covar}_{ij} - \mathbf{covar}^{\mathrm{h}}_{ij})^2}{P^{\mathrm{h}}_{ij}}$$

Deze heeft een lokaal minimum 0 bij $e_{ij} = 0$. De termen e_{ij} moeten ook voldoen aan de somregel voor i en j apart omdat het verschillen van covarianties zijn. Dat is dezelfde beperking als in het onafhankelijke geval. Hierdoor is chi-kwadraat als een kwadratische vorm op te vatten in de excesvariabelen e_{ij}:

$$\begin{aligned}
\chi^2_j &= \sum_i \chi^2_{ij} = N\sum_i \frac{e^2_{ij}}{P^{\mathrm{h}}_{ij}} = N{\sum_i}' \frac{e^2_{ij}}{P^{\mathrm{h}}_{ij}} + N\frac{e^2_{kj}}{P^{\mathrm{h}}_{kj}} \\
&= N{\sum_i}' \frac{e^2_{ij}}{P^{\mathrm{h}}_{ij}} + N\frac{(-\sum'_i e_{ij})^2}{P^{\mathrm{h}}_{kj}} = N{\sum_i}'{\sum_{i'}}' e_{ij}\left(\delta_{i,i'}\frac{1}{P^{\mathrm{h}}_{ij}} + \frac{1}{P^{\mathrm{h}}_{kj}}\right) e_{i'j}
\end{aligned}$$

Merk op, dat het belangrijkste verschil met het onafhankelijke geval is, dat de symmetrische matrix nu van j (de gevolgen) afhangt. Diagonalisatie zou even zo goed kunnen, maar de eigenwaarden en eigenvectoren zouden afhangen van j. We zouden dientengevolge geen 'onafhankelijke oorzaken λ' krijgen, omdat iedere gevolg j zijn eigen λ's heeft. De afhankelijkheid heeft oorzaken en gevolgen als het ware 'gemengd', zodat we ze als totaal moeten opvatten: we moeten naast de oorzaken i ook de gevolgen j meenemen.

Bij het onafhankelijke geval hadden we apart voor de oorzaken de dimensie $|i|$ en voor de gevolgen dimensie $|j|$. Bij de als koppel genomen variabelen 'verveelvoudigt' de dimensie tot $|ij| = (k-1)(l-1)$. Dat gaat als volgt: in chi-kwadraat splitsen we de som over j kunstmatig tot sommaties over j (bij i), of j' (bij i'), en l:

$$\chi^2/N = \sum_j \chi_j^2/N = \sum_j{}' \sum_i{}' \frac{e_{ij}^2}{P_{ij}^{\mathrm{h}}} + \sum_i{}' \frac{e_{il}^2}{P_{il}^{\mathrm{h}}} + \sum_j{}' \frac{e_{kj}^2}{P_{kj}^{\mathrm{h}}} + \frac{e_{kl}^2}{P_{kl}^{\mathrm{h}}}$$

$$e_i = \Sigma_j{}' e_{i,j} = -e_{il} \quad e_j = \Sigma_i{}' e_{i,j} = -e_{kj} \quad e = \Sigma_i{}' \Sigma_j{}' e_{i,j} = e_{kl}$$

$$= \sum_j{}' \sum_i{}' \frac{e_{ij}^2}{P_{ij}^{\mathrm{h}}} + \sum_i{}' \frac{(-\sum_j{}' e_{ij})^2}{P_{il}^{\mathrm{h}}} + \sum_j{}' \frac{(-\sum_i' e_{ij})^2}{P_{kj}^{\mathrm{h}}} + \frac{(\sum_i' \sum_j' e_{ij})^2}{P_{kl}^{\mathrm{h}}}$$

$$= \sum_{i,i'}{}' \sum_{j,j'}{}' e_{ij} F_{ij,i'j'} e_{i'j'}$$

$$\boxed{F_{ij,i'j'} = \left(\delta_{j,j'} \left(\delta_{i,i'} \frac{1}{P_{i'j'}^{\mathrm{h}}} + \frac{1}{P_{kj'}^{\mathrm{h}}}\right) + \left(\delta_{i,i'} \frac{1}{P_{i'l}^{\mathrm{h}}} + \frac{1}{P_{kl}^{\mathrm{h}}}\right)\right)}$$

Om de gedachte te bepalen, bijvoorbeeld $k = l = 2$, $i = j = 1$: De χ^2-matrix bestaat slechts uit het getal $1/P_{ij}^{\mathrm{h}} + 1/P_{il}^{\mathrm{h}} + 1/P_{kj}^{\mathrm{h}} + 1/P_{kl}^{\mathrm{h}}$, de som van alle omgekeerde kansen. Dat komt bekend voor! Ook met gecombineerde ij-index is de matrix symmetrisch vanwege de deltafuncties.

3.1 De eigenwaarden van de matrix F van χ^2

2019: We bewijzen dat F alleen positieve eigenwaarden heeft, dus geen nulruimte heeft, omdat de matrix *positief-definiet* is. Daartoe tonen we aan dat voor iedere vector $z_{ij} \neq 0$ de kwadratische vorm met χ^2 positief is (alleen 0 als $z = 0$):

$$\Sigma'_{ij} \Sigma'_{i'j'} z_{ij} F_{ij,i'j'} z_{i'j'}$$

$$= \Sigma'_{ij} \Sigma'_{i'j'} z_{ij} \left(\delta_{i,i'} \delta_{j,j'} \frac{1}{P_{ij}^{\mathrm{h}}} + \delta_{j,j'} \frac{1}{P_{kj}^{\mathrm{h}}} + \delta_{i,i'} \frac{1}{P_{il}^{\mathrm{h}}} + \frac{1}{P_{kl}^{\mathrm{h}}}\right) z_{i'j'}$$

$$= \Sigma'_{ij} \frac{z_{ij}^2}{P_{ij}^{\mathrm{h}}} + \Sigma'_i \frac{(\Sigma'_j z_{ij})^2}{P_{il}^{\mathrm{h}}} + \Sigma'_j \frac{(\Sigma'_i z_{ij})^2}{P_{kj}^{\mathrm{h}}} + \frac{(\Sigma'_{ij} z_{ij})^2}{P_{kl}^{\mathrm{h}}} \geq 0$$

Het rechterlid bestaat uit een som van positieve breuken—alle kansen in de noemer zijn positief en alle tellers zijn kwadraten—dus de som is alleen dan 0 als alle breuken apart 0 zijn. In de eerste term is nul als alle breuken nul: als alle $z_{ij} = 0$ zijn. Daardoor worden alle andere termen ook nul. Daarmee is een eigenwaarde nul uitgesloten, en blijken alle eigenwaarden positief te zijn. Door de specificatie $z_{ij} = v_{ij\lambda}$, de genormalizeerde eigenvector met eigenwaarde $1/\lambda$ en $\Sigma'_{ij} v_{ij\lambda}^2 = 1$, krijgen we

$$\Sigma'_{ij} \Sigma'_{i'j'} v_{ij} F_{ij,i'j'} v_{i'j'\lambda} = \Sigma'_{ij} v_{ij\lambda} (1/\lambda) v_{ij\lambda} = 1/\lambda > \Sigma'_{ij} \frac{v_{ij\lambda}^2}{P_{ij}^{\mathrm{h}}} \geq 1/\mathbf{max}(P_{ij}^{\mathrm{h}})$$

We concluderen dat de eigenwaarde ook naar boven is begrensd: $0 < \lambda < \mathbf{max}(P_{ij}^{\mathrm{h}}) < 1$.
Er is nog een manier om naar de eigenwaarden te kijken: de som van alle eigenwaarden is gelijk aan de de som van de diagonaalelementen:

$$\Sigma_\lambda (1/\lambda) = \Sigma'_{ij} F_{ij,ij} = \Sigma'_{ij} (1/P_{ij}^{\mathrm{h}} + 1/P_{il}^{\mathrm{h}} + 1/P_{kj}^{\mathrm{h}} + 1/P_{kl}^{\mathrm{h}})$$

Bedenk dat het aantal eigenwaarden $n = (k-1)(l-1)$ gelijk is aan het aantal kansen, zodat de gemiddelde omgekeerde eigenwaarde gelijk is aan de som van de gemiddelde omgekeerde kansen:

$$\overline{1/\lambda} = \overline{1/P_{ij}^{\mathrm{h}}} + \overline{1/P_{il}^{\mathrm{h}}} + \overline{1/P_{kj}^{\mathrm{h}}} + 1/P_{kl}^{\mathrm{h}}$$

Dus is enerzijds en anderzijds:

$$1/\mathbf{max}(P_{ij}^{\mathrm{h}}) < 1/\mathbf{max}(\lambda) \leq \overline{1/\lambda} = \overline{1/P_{ij}^{\mathrm{h}}} + \overline{1/P_{il}^{\mathrm{h}}} + \overline{1/P_{kj}^{\mathrm{h}}} + 1/P_{kl}^{\mathrm{h}} \leq 1/\mathbf{min}(\lambda)$$

In overeenstemming daarmee veronderstellen we dat, net zoals bij de onafhankelijke variabelen, de eigenwaarden λ steeds tussen opeenvolgende P_{ij}^{h} zullen liggen. Dat is echter niet bewezen.

3.2 Het minimum van χ^2

2019: De covarianties γ worden bepaald door de eis dat χ^2 zo klein als mogelijk is. Met de beperkte indices $i = 1 \cdots k-1$ en $j = 1 \cdots l-1$, hebben we $(k-1)(l-1)$ onafhankelijke variabelen $\gamma_{ij} = \mathbf{covar}_{ij}$, die we zodanig kunnnen variëren dat χ^2 minimaal wordt. De variabele $\gamma_{i,j}$ zit, zoals we zagen, in $e_{i,j}$ maar is ook verborgen aanwezig in e_{kj}, e_{il} en e_{kl}, omdat de som van de covarianties nul is:

$$\begin{aligned}
\chi^2/N &= \sum_j{}' \sum_i{}' \frac{e_{ij}^2}{P_{ij}^{\mathrm{h}}} + \sum_i{}' \frac{e_{il}^2}{P_{il}^{\mathrm{h}}} + \sum_j{}' \frac{e_{kj}^2}{P_{kj}^{\mathrm{h}}} + \frac{e_{kl}^2}{P_{kl}^{\mathrm{h}}} \\
e_{i,j} &= P_{ij} - P_{ij}^{\mathrm{h}} = \gamma_{ij} - \gamma_{ij}^{\mathrm{h}} \\
(e_{i,j})' &= \delta_i \delta_j \quad (e_{il})' = -\delta_i \quad (e_{kj})' = -\delta_j \quad (e_{kl})' = 1
\end{aligned}$$

waardoor het verdwijnen van de afgeleide kan worden geformuleerd als:

$$\begin{aligned}
0 &= \sum_j{}' \sum_i{}' (\frac{e_{ij}^2}{P_{ij}^{\mathrm{h}}})' + \sum_i{}' (\frac{e_{il}^2}{P_{il}^{\mathrm{h}}})' + \sum_j{}' (\frac{e_{kj}^2}{P_{kj}^{\mathrm{h}}})' + (\frac{e_{kl}^2}{P_{kl}^{\mathrm{h}}})' \\
0 &= \frac{e_{ij}}{P_{ij}^{\mathrm{h}}} - \frac{e_{il}}{P_{il}^{\mathrm{h}}} - \frac{e_{kj}}{P_{kj}^{\mathrm{h}}} + \frac{e_{kl}}{P_{kl}^{\mathrm{h}}} \\
0 &= \frac{e_{ij}}{P_{ij}^{\mathrm{h}}} + \frac{e_i}{P_{il}^{\mathrm{h}}} + \frac{e_j}{P_{kj}^{\mathrm{h}}} + \frac{e}{P_{kl}^{\mathrm{h}}} \\
0 &= \sum_{i'}{}' \sum_{j'}{}' \left(\delta_{j,j'} \left(\delta_{i,i'} \frac{1}{P_{i'j'}^{\mathrm{h}}} + \frac{1}{P_{kj'}^{\mathrm{h}}} \right) + \left(\delta_{i,i'} \frac{1}{P_{i'l}^{\mathrm{h}}} + \frac{1}{P_{kl}^{\mathrm{h}}} \right) \right) e_{i'j'} \\
&\Sigma'_{i'j'} F_{ij,i'j'} e_{i'j'} = 0
\end{aligned}$$

Er zijn $(k-1)(l-1)$ vergelijkingen, van de eerste orde, voor evenzoveel variabelen e_{ij}, of γ_{ij}. Die hebben ook een e=0 oplossing, want als alle onafhankelijke e_{ij}=0 zijn, dan zijn, vanwege de somnulregel, ook de marginale e_{il}=0, e_{kj}=0 en e_{kl}=0. Verder voldoet iedere eigenvector met eigenwaarde nul van de χ^2-matrix aan de evenwichtsvoorwaarde. We kunnen aantonen dat er geen nul eigenwaarde, en dus geen nulvector, is: $e = 0$ is het enige minimum.

Houden we rekening met de variabele $P_{ij}^{\mathrm{h}} = P_{ij} - e_{ij}$ in de teller van χ^2, dan wordt de minimum voorwaarde iets ingewikkelder:

$$\begin{aligned}
(\chi^2)' &= \sum_j{}' \sum_i{}' (\frac{e_{ij}^2}{P_{ij}^{\mathrm{h}}})' + \sum_i{}' (\frac{e_{il}^2}{P_{il}^{\mathrm{h}}})' + \sum_j{}' (\frac{e_{kj}^2}{P_{kj}^{\mathrm{h}}})' + (\frac{e_{kl}^2}{P_{kl}^{\mathrm{h}}})' \\
&= 2\frac{e_{ij}}{P_{ij}^{\mathrm{h}}} + (\frac{e_{ij}}{P_{ij}^{\mathrm{h}}})^2 - 2\frac{e_{il}}{P_{il}^{\mathrm{h}}} - (\frac{e_{il}}{P_{il}^{\mathrm{h}}})^2 - 2\frac{e_{kj}}{P_{kj}^{\mathrm{h}}} - (\frac{e_{kj}}{P_{kj}^{\mathrm{h}}})^2 + 2\frac{e_{kl}}{P_{kl}^{\mathrm{h}}} + (\frac{e_{kl}}{P_{kl}^{\mathrm{h}}})^2 \\
&= (\frac{e_{ij}}{P_{ij}^{\mathrm{h}}}+1)^2 - (\frac{e_{il}}{P_{il}^{\mathrm{h}}}+1)^2 - (\frac{e_{kj}}{P_{kj}^{\mathrm{h}}}+1)^2 + (\frac{e}{P_{kl}^{\mathrm{h}}}+1)^2 \\
&= \boxed{(\frac{P_{ij}}{P_{ij}^{\mathrm{h}}})^2 - (\frac{P_{il}}{P_{il}^{\mathrm{h}}})^2 - (\frac{P_{kj}}{P_{kj}^{\mathrm{h}}})^2 + (\frac{P_{kl}}{P_{kl}^{\mathrm{h}}})^2 = 0}
\end{aligned}$$

Ook nu is er een e=0 oplossing, als $P_{ij}^{\mathrm{h}} = P_{ij}$, maar vanwege de verschillende plus en min bijdragen is ook een $e \neq 0$ oplossing denkbaar. Echter, de tweede afgeleide naar $e_{ij}$is altijd positief:

$$\begin{aligned}
(\chi^2)'' &= ((\frac{P_{ij}}{P_{ij}^{\mathrm{h}}})^2)' - ((\frac{P_{il}}{P_{il}^{\mathrm{h}}})^2)' - ((\frac{P_{kj}}{P_{kj}^{\mathrm{h}}})^2)' + ((\frac{P_{kl}}{P_{kl}^{\mathrm{h}}})^2)' \\
&= +2\frac{P_{ij}^2}{{P_{ij}^{\mathrm{h}}}^3} + 2\frac{P_{il}^2}{{P_{il}^{\mathrm{h}}}^3} + 2\frac{P_{kj}^2}{{P_{kj}^{\mathrm{h}}}^3} + 2\frac{P_{kl}^2}{{P_{kl}^{\mathrm{h}}}^3} > 0
\end{aligned}$$

Ook de 'algemene' tweede afgeleide in willekeurige richting $(\Sigma'_{ij}\lambda_{ij}\delta e_{ij})^2$, met $\Sigma'_{ij}\lambda_{ij}=1$, is een som van positieve termen, zodat een niet nul oplossing $e\neq 0$ is uitgesloten:

$$
\begin{aligned}
&(\Sigma'_{ij}\lambda_{ij}\delta e_{ij})(\Sigma'_{i'j'}\lambda_{i'j'}\delta e_{i'j'})\left[\sum_{i''j''}{}'\frac{e^2_{i''j''}}{P^{\mathrm{h}}_{i''j''}}+\sum_{i''}{}'\frac{e^2_{i''l}}{P^{\mathrm{h}}_{i''l}}+\sum_{j''}{}'\frac{e^2_{kj''}}{P^{\mathrm{h}}_{kj''}}+\frac{e^2_{kl}}{P^{\mathrm{h}}_{kl}}\right]=\\
=&(\Sigma'_{ij}\lambda_{ij}\delta e_{ij})(\Sigma'_{i'j'}\lambda_{i'j'}\left[\frac{P^2_{i'j'}}{{P^{\mathrm{h}}_{i'j'}}^2}-\frac{P^2_{i'l}}{{P^{\mathrm{h}}_{i'l}}^2}-\frac{P^2_{kj'}}{{P^{\mathrm{h}}_{kj'}}^2}+\frac{P^2_{kl}}{{P^{\mathrm{h}}_{kl}}^2}\right]\\
=&2(\Sigma'_{ij}\lambda_{ij})(\Sigma'_{i'j'}\lambda_{i'j'}\left[\delta_{ij,i'j'}\frac{P^2_{ij}}{{P^{\mathrm{h}}_{ij}}^3}+\delta_{ii'}\frac{P^2_{il}}{{P^{\mathrm{h}}_{il}}^3}+\delta_{jj'}\frac{P^2_{kj}}{{P^{\mathrm{h}}_{kj}}^3}+\frac{P^2_{kl}}{{P^{\mathrm{h}}_{kl}}^3}\right]\\
=&2\Sigma'_{ij}\lambda^2_{ij}\frac{P^2_{ij}}{{P^{\mathrm{h}}_{ij}}^3}+2\Sigma'_i(\Sigma'_j\lambda_{ij})^2\frac{P^2_{il}}{{P^{\mathrm{h}}_{il}}^3}+2\Sigma'_j(\Sigma'_i\lambda_{ij})^2\frac{P^2_{kj}}{{P^{\mathrm{h}}_{kj}}^3}+2(\Sigma'_{ij}\lambda_{ij})^2\frac{P^2_{kl}}{{P^{\mathrm{h}}_{kl}}^3}>0
\end{aligned}
$$

3.3 Eigenwaarde- en eigenvector-vergelijkingen en deelsom variabelen $v_{i\lambda}$, $v_{j\lambda}$ en v_λ.

Voor het minimaliseren van χ^2 zoeken we de eigenwaarden en eigenvectoren van de F-matrix. De omgekeerde 'eigenwaarden' λ (voor oorzaak èn gevolg samen) en eigenvectoren $v_{ij\lambda}$ moeten ook nu aan eigenwaarde-vergelijkingen voldoen:

$$\sum_{i'}{}'\sum_{j'}{}'\left(\delta_{j,j'}\left(\delta_{i,i'}\frac{1}{P^{\mathrm{h}}_{i'j'}}+\frac{1}{P^{\mathrm{h}}_{kj'}}\right)+\left(\delta_{i,i'}\frac{1}{P^{\mathrm{h}}_{i'l}}+\frac{1}{P^{\mathrm{h}}_{kl}}\right)\right)v_{i'j'\lambda}=v_{ij\lambda}/\lambda$$

Formeel kunnen we chi-kwadraat—netzo als in het onafhankelijke geval, maar nu met oorzaak en gevolg tezamen—schrijven in termen van de λ-excesvariabelen, dus in de λ-covarianties:

$$
\begin{aligned}
e_\lambda=&\sum_{ij}{}'v_{ij\lambda}e_{ij}=\sum_{ij}{}'v_{ij\lambda}(P_{ij}-P^{\mathrm{h}}_{ij})=\sum_{ij}{}'v_{ij\lambda}(\gamma_{ij}-\gamma^{\mathrm{h}}_{ij})\\
e_\lambda=&P_\lambda-P^{\mathrm{h}}_\lambda=\gamma_{ij}-\gamma^{\mathrm{h}}_{ij}=\mathbf{covar}_\lambda-\mathbf{covar}^h_\lambda=S_\lambda(r_\lambda-\rho)
\end{aligned}
$$

en in de $(k-1)(l-1)$ eigenwaarden λ:

$$\boxed{\chi^2=N\sum_\lambda{}'\frac{e^2_\lambda}{\lambda}=N\sum_\lambda{}'\frac{S^2_\lambda}{\lambda}(r_\lambda-\rho_\lambda)^2}$$

Voor het kunnen oplossen van de eigenvectorvergelijking definiëren we ook nu *(deel)somvariabelen*:

$$v_{j\lambda}=\Sigma'_i v_{ij\lambda}\quad v_{i\lambda}=\Sigma'_j v_{ij\lambda}\quad \Sigma''_{ij}v_{ij\lambda}=\Sigma'_i v_{i\lambda}=\Sigma'_j v_{j\lambda}=v_\lambda$$

De laatste vergelijkingen zijn de *consistentievergelijkingen* voor $v_{i\lambda}$ en $v_{j\lambda}$. De eigenvector, in termen van deelsomvariabelen en eigenwaarde, is:

$$\boxed{v_{ij\lambda}=A_{\lambda ij}\left(\frac{v_{j\lambda}}{P^{\mathrm{h}}_{kj}}+\frac{v_{i\lambda}}{P^{\mathrm{h}}_{il}}+\frac{v_\lambda}{P^{\mathrm{h}}_{kl}}\right)}\quad\boxed{A_{\lambda ij}=\frac{1}{1/\lambda-1/P^{\mathrm{h}}_{ij}}}$$

De $k+l-1$ variabelen v_i, v_j, v bepalen de $(k-1)(l-1)$ variabelen v_{ij}. Bijgevolg moet de dimensie van de nulruimte van $A_{\lambda ij}$ minstens gelijk $(k-1)(l-1)-(k+l-1)=(k-2)(l-2)-2\geq 0$, dus $k,l>3$. Voor grotere aantallen variabelen wordt de nulruimte al snel relatief groot: $((k-1)(l-1)-(k-1)-(l-1))/(k-1)(l-1)=1-1/(k-1)-1/(l-1)$; bijvoorbeeld met $k,l=5$ is dat 50%, voor $k,l=9$ is 75%..

Vanwege de orthonormaliteit van de $v_{ij\lambda}$, kunnen we omgekeerd de excesvariabelen e_{ij} uitdrukken in de gemengde excesvariabelen e_λ:

$$e_{ij}=\sum_\lambda{}'v_{ij\lambda}e_\lambda\quad e_i=-e_{il}=\sum_\lambda{}'v_{i\lambda}e_\lambda\quad e_j=-e_{kj}=\sum_\lambda{}'v_{j\lambda}e_\lambda\quad e=-e_{kl}=\sum_\lambda{}'v_\lambda e_\lambda$$

Het accent bij de sommatie herinnert er aan dat er net zoveel λ waarden zijn als er ij waarden zijn. Voor χ^2 vinden we, omgekeerd, met e_{ij} uitgedrukt in e_λ:

$$\begin{aligned}
\chi^2/N &= \sum_{ij}{}' \frac{e_{ij}^2}{P_{ij}^{\mathrm{h}}} + \sum_i{}' \frac{e_{il}^2}{P_{il}^{\mathrm{h}}} + \sum_j{}' \frac{e_{kj}^2}{P_{kj}^{\mathrm{h}}} + \frac{e_{kl}^2}{P_{kl}^{\mathrm{h}}} \\
&= \sum_{ij}{}' \frac{(\sum_\lambda{}' v_{ij\lambda} e_\lambda)^2}{P_{ij}^{\mathrm{h}}} + \sum_i{}' \frac{(\sum_\lambda{}' v_{i\lambda} e_\lambda)^2}{P_{il}^{\mathrm{h}}} + \sum_j{}' \frac{(\sum_\lambda{}' v_{j\lambda} e_\lambda)^2}{P_{kj}^{\mathrm{h}}} + \frac{(\sum_\lambda{}' v_\lambda e_\lambda)^2}{P_{kl}^{\mathrm{h}}} \\
&= \sum_\lambda{}' \sum_{\lambda'}{}' e_\lambda e_{\lambda'} \sum_{ij}{}' v_{ij\lambda} \sum_{i'j'}{}' \left(\frac{\delta_{ij,i'j'}}{P_{ij}^{\mathrm{h}}} + \frac{\delta_{i,i'}}{P_{il}^{\mathrm{h}}} + \frac{\delta_{j,j'}}{P_{kj}^{\mathrm{h}}} + \frac{1}{P_{kl}^{\mathrm{h}}} \right) v_{i'j'\lambda'} \\
&= \sum_\lambda{}' \sum_{\lambda'}{}' e_\lambda e_{\lambda'} \sum_{ij}{}' v_{ij\lambda} \frac{v_{i'j'\lambda'}}{\lambda'} = \sum_\lambda{}' \frac{e_\lambda^2}{\lambda}
\end{aligned}$$

3.3.1 'Oplossing' met geschaalde variabelen $x_{i\lambda}$, $x_{j\lambda}$ en x_λ

Verdere bewerkingen worden iets doorzichtiger als de v-variabelen worden geschaald met de 'marge'kans tot $x_{i\lambda}$, $x_{j\lambda}$ en x_λ; benoem verder de factor $A_{ij\lambda}$:

$$\begin{aligned}
x_{i\lambda} &= v_{i\lambda}/P_{il}^{\mathrm{h}} \quad x_{j\lambda} = v_{j\lambda}/P_{kj}^{\mathrm{h}} \quad x_\lambda = v_\lambda/P_{kl}^{\mathrm{h}} \\
A_{ij\lambda} &= \frac{1}{1/\lambda - 1/P_{ij}^{\mathrm{h}}} \quad v_{ij\lambda} = A_{ij\lambda}(x_{i\lambda} + x_{j\lambda} + x_\lambda)
\end{aligned}$$

De 'bekenden' v_i, v_j en v zijn deelsommen van v_{ij}:

$$\begin{aligned}
v_{i\lambda} &= \Sigma_j' v_{ij\lambda} \leftrightarrow P_{il}^{\mathrm{h}} x_{i\lambda} = A_{i\lambda}(x_{i\lambda} + x_\lambda) + \Sigma_j' A_{ij\lambda} x_{j\lambda} \quad A_{i\lambda} = \Sigma_j' A_{ij\lambda} \\
v_{j\lambda} &= \Sigma_i' v_{ij\lambda} \leftrightarrow P_{kj}^{\mathrm{h}} x_{j\lambda} = A_{j\lambda}(x_{j\lambda} + x_\lambda) + \Sigma_j' A_{ij\lambda} x_{i\lambda} \quad A_{j\lambda} = \Sigma_i' A_{ij\lambda} \\
v_\lambda &= \Sigma_i' \Sigma_j' v_{ij\lambda} = \Sigma_i' v_{i\lambda} = \Sigma_j' v_{j\lambda} \leftrightarrow \Sigma_i' P_{il}^{\mathrm{h}} x_{i\lambda} = \Sigma_j' P_{kj}^{\mathrm{h}} x_{j\lambda} = \\
&\quad P_{kl}^{\mathrm{h}} x_\lambda = \Sigma_i' A_{i\lambda} x_{i\lambda} + \Sigma_j' A_{j\lambda} x_{j\lambda} + A_\lambda x_\lambda \quad A_\lambda = \Sigma_i' A_{i\lambda} = \Sigma_j' A_{j\lambda} = \Sigma_i' \Sigma_j' A_{ij\lambda}
\end{aligned}$$

De eerste $k{+}l{-}2$ vergelijkingen drukken $k{+}l{-}2$ variabelen x_i en x_j in x uit (**ter vereenvouding** van de schrijfwijze **wordt λ weggelaten**). In matrix-vector notatie:

$$\begin{aligned}
B_i x_i - \Sigma_j' A_{ij} x_j &= A_i x \quad B_i = P_{il}^{\mathrm{h}} - A_i \\
-\Sigma_i' A_{ij} x_i + B_j x_j &= A_j x \quad B_j = P_{kj}^{\mathrm{h}} - A_j
\end{aligned}$$

De laatste volgende 3 vergelijkingen zijn consistentievergelijkingen waarmee alle variabelen in x worden uitgedrukt, dus ook $v_{ij\lambda}$. Uiteindelijk wordt x bepaald door de normering van $v_{ij\lambda}$.

$$\begin{aligned}
\Sigma_i' A_i x_i + \Sigma_j' A_j x_j &= Bx \qquad B = P_{kl}^{\mathrm{h}} - A \\
\Sigma_i' P_{il}^{\mathrm{h}} x_i &= P_{kl}^{\mathrm{h}} x \\
\Sigma_j' P_{kj}^{\mathrm{h}} x_j &= P_{kl}^{\mathrm{h}} x
\end{aligned}$$

De tweede consistentie vergelijking, met P_{il}^{h}, geeft samen met de $k{-}1$ i-vergelijkingen, de eerste consistentievergelijking, met A_i, A_j en B. Evenzo geeft de derde consistentie vergelijking, met P_{kj}^{h}, samen met de $l{-}1$ j-vergelijkingen de eerste consistentie vergelijking. Deze vergelijkingen zijn dus, tezamen met de vorige vergelijkingen, afhankelijk, zodat er slechts één onafhankelijke consistentievergelijking is.

Vat B_i op als diagonale matrix en x_i, P_{il}^{h} als vectoren in de i-ruimt. Evenzo duaal B_j en x_j en P_{kj}^{h} in de j-ruimte. Dan is A_{ij} een rechthoekige matrix tussen de i en j ruimte, met $A_i = \Sigma_j' A_{ij} = Au_j$ een vector in de i-ruimte en $A_j = \Sigma_i' A_{ij} = A^t u_i$ een vector in de j-ruimte. Daarbij zijn u_i en u_j vectoren in de i en j-ruimte met coëfficienten 1. De vergelijkingen zijn:

$$\begin{pmatrix} B_i & -A \\ -A^t & B_j \end{pmatrix} \begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} Au_j \\ A^t u_i \end{pmatrix} x$$

$$u_j^t A^t x_i + u_i^t A x_j = Bx \qquad {P_{il}^{\mathrm{h}}}^t x_i = P_{kl}^{\mathrm{h}} x \qquad {P_{kj}^{\mathrm{h}}}^t x_j = P_{kl}^{\mathrm{h}} x$$

Zijn B_i en B_j onafhankelijk, maar ook de Schur-complementen $\overline{B_i}$ en $\overline{B_j}$, dan lossen we x_i en x_j op in x:

$$\overline{B_i} = B_i - AB_j^{-1}A^t \quad \overline{B_j} = B_j - A^tB_i^{-1}A$$

$$\begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} B_i & -A \\ -A^t & B_j \end{pmatrix}^{-1} \begin{pmatrix} Au_j \\ A^tu_i \end{pmatrix} x = \begin{pmatrix} \overline{B_i}^{-1} & \overline{B_i}^{-1}AB_j^{-1} \\ \overline{B_j}^{-1}A^tB_i^{-1} & \overline{B_j}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} Au_j \\ A^tu_i \end{pmatrix} x$$

$$= \begin{pmatrix} \overline{B_i}^{-1}(Au_j+AB_j^{-1}A^tu_i) \\ \overline{B_j}^{-1}(A^tu_i+A^tB_i^{-1}Au_j) \end{pmatrix} x$$

Met de speciale eigenschap $B_j^{-1}A_j{=}B_j^{-1}P^{\mathrm{h}}_{kj}{-}u_j$ (en duaal), wordt de oplossing herschreven:

$$Au_j+A(B_j^{-1}A^tu_i)) = Au_j+A(B_j^{-1}P^{\mathrm{h}}_{kj}-u_j) = AB_j^{-1}P^{\mathrm{h}}_{kj}$$

De oplossing wordt daarmee:

$$\boxed{x_i = \overline{B_i}^{-1}AB_j^{-1}P^{\mathrm{h}}_{kj}x} \qquad \boxed{x_j = \overline{B_j}^{-1}A^tB_i^{-1}P^{\mathrm{h}}_{il}x}$$

Deze oplossing moet voldoen aan de consistentievergelijkingen, onafhankelijk van x:

$$\boxed{P^{\mathrm{h}}_{kl} = {P^{\mathrm{h}}_{il}}^t\overline{B_i}^{-1}AB_j^{-1}P^{\mathrm{h}}_{kj}} \quad \boxed{P^{\mathrm{h}}_{kl} = {P^{\mathrm{h}}_{kj}}^t\overline{B_j}^{-1}A^tB_i^{-1}P^{\mathrm{h}}_{il}}$$

De algemene consistentievergelijking is:

$$B = P^{\mathrm{h}}_{kl} - u_i^tAu_j = u_j^tA^t\overline{B_i}^{-1}AB_j^{-1}P^{\mathrm{h}}_{kj} + u_i^tA\overline{B_j}^{-1}A^tB_i^{-1}P^{\mathrm{h}}_{il}$$

De x_j consistentievergelijking volgt uit die van x_i (of omgekeerd), niet alleen met dualiteit, maar ook rechtstreeks met algemene basiseigenschappen van $\overline{B}$ matrices [5]

$$\overline{B_i}^{-1}AB_j^{-1} = B_i^{-1}A\overline{B_j}^{-1} \qquad B_j^{-1}A^t\overline{B_i}^{-1} = \overline{B_j}^{-1}A^tB_i^{-1}$$

2019: Verrassenderwijs is er een oplossing x_i (x_j), met bijbehorende consistentievergelijking, door niet $A_j/B_j = P^{\mathrm{h}}_{kj}/B_j - 1$, maar alternatief $A_i = P^{\mathrm{h}}_{il} - B_i$, toe te passen:

$$\begin{aligned} x_i/x &= \overline{B_i}^{-1}(Au_j+AB_j^{-1}A^tu_i) = \overline{B_i}^{-1}((P^{\mathrm{h}}_{il}-B_iu_i)+AB_j^{-1}A^tu_i) \\ &= \overline{B_i}^{-1}P^{\mathrm{h}}_{il} - \overline{B_i}^{-1}(B_i-AB_j^{-1}A^t)u_i \\ &= \overline{B_i}^{-1}P^{\mathrm{h}}_{il} - \overline{B_i}^{-1}\overline{B_i}u_i \end{aligned}$$

$$\boxed{x_i/x = \overline{B_i}^{-1}P^{\mathrm{h}}_{il} - u_i} \qquad \boxed{x_j/x = \overline{B_j}^{-1}P^{\mathrm{h}}_{kj} - u_j}$$

$$\rightarrow P^{\mathrm{h}}_{kl} = {P^{\mathrm{h}}_{il}}^t\overline{B_i}^{-1}P^{\mathrm{h}}_{il} - {P^{\mathrm{h}}_{il}}^tu_i = {P^{\mathrm{h}}_{il}}^t\overline{B_i}^{-1}P^{\mathrm{h}}_{il} - P_l + P^{\mathrm{h}}_{kl}$$

$$\boxed{P_l = {P^{\mathrm{h}}_{il}}^t\overline{B_i}^{-1}P^{\mathrm{h}}_{il}} \qquad \boxed{P_k = {P^{\mathrm{h}}_{kj}}^t\overline{B_j}^{-1}P^{\mathrm{h}}_{kj}}$$

$$P^{\mathrm{h}}_{kl} + u_i^tAu_j = u_j^tA^t\overline{B_i}^{-1}P^{\mathrm{h}}_{il} + u_i^tA\overline{B_j}^{-1}P^{\mathrm{h}}_{kj}$$

Deze consistentievergelijkingen zijn alle gelijkwaardig. De x_j consistentie volgt uit die van x_i, en omgekeerd; de x_i consistentie volgt uit de laatste consistentie vergelijking. De oplossingen λ van deze consistentievergelijking moeten de $(k-1)(l-1)$ eigenwaarden van eigenvector v zijn. Het probleem is echter het omkeren van $\overline{B}$. Is dat wel mogelijk zolang λ onbekend is?

[5] op Wikipedia zijn bijvoorbeeld de eigenschappen van deze 'Schur-complement'en te vinden; zie ook de appendix: Eigenschappen van Schur-complementen.

3.3.2 Oplossing met D_i- of D_j-matrices.

2015: Uitgangspunt voor bepaling van eigenwaarden en eigenvectoren is de chi-kwadraatmatrix F voor de samengevoegde variabelen. Karakteristiek van deze matrix is het sterke diagonale karakter. Dat blijft behouden als we de chi-kwadraatmatrix F verminderen met $1/\lambda$ maal de eenheidsmatrix vanwege de eigenvectorbepaling waardoor matrix E ontstaat. We beschouwen de χ^2-matrix F als opgebouwd uit $j = 1 \ldots l{-}1$ bij $l{-}1$ submatrices van $i = 1 \ldots k{-}1$ bij $k{-}1$. Op de diagonaal van de matrix komt de j-afhankelijke ($j = 1 \cdots l-1$) submatrix D_j, daarnaast komt in ieder cel de j-onafhankelijke submatrix D_l:

$$F_{j,j'} = \delta_{jj'} D_j + D_l$$

$$D_{j\,ii'} = \delta_{i,i'} \frac{1}{P^{\mathrm{h}}_{ij}} + \frac{1}{P^{\mathrm{h}}_{kj}} \qquad D_{l\,ii'} = \delta_{i,i'} \frac{1}{P^{\mathrm{h}}_{il}} + \frac{1}{P^{\mathrm{h}}_{kl}}$$

$$D_{j\lambda} = D_j - \delta_{ii'} \frac{1}{\lambda} \qquad E_{j,j'} = \delta_{jj'} D_{j\lambda} + D_l$$

De matrices F, D_l, D_j, $D_{j\lambda}$ en E hebben alle dezelfde structuur: opgebouwd uit niet-diagonalen elementen, die alle gelijk D_l (i-matrix) of n (scalar) zijn, en diagonale elementen $D_j{+}D_l$ (i-matrix) of $d_i{+}n$ (scalar). Dit type matrices noemen we symbolisch $F = \delta_{jj'} D_j{+}D_l$ (eventueel is de D_j dus $D_{j\lambda}$ als F is E).
Eigenwaarde en eigenvector van F vinden we als nulruimte van E:

$$0 = \Sigma'_{j'} E_{jj'} V j' = \Sigma'_{j'} (\delta_{jj'} D_{j\lambda}{+}D_l) V_{j'} = D_{j\lambda} V_j + D_l \Sigma'_j V_j$$

$$D_{j\lambda} V_j = -D_l \Sigma'_j V_j = -D_l v_{i\lambda} = -y_{\lambda i} \quad v_{i\lambda} = \Sigma'_j V_j = D_l^{-1} y_{\lambda i}$$

We zullen zien dat D_l inderdaad omkeerbaar is, zodat $v_{i\lambda}$ uitdrukbaar is in $y_{\lambda i}$; $y_{\lambda i}$ hangt alleen af van λ!
Als $y_{\lambda i} = 0$ zou zijn, dan is $1/\lambda$ voor alle j eigenwaarde van D_j, en voldoen de eigenvectoren V_j aan de somnulconditie:

$$D_j V_j = V_j/\lambda \quad \Sigma'_j V_j = 0$$

Het is hoogst onwaarschijnlijk dat alle D_j een gemeenschappelijke eigenwaarde zouden hebben. Wij zullen dus aannemen dat de eigenwaarden van de D_j geen eigenwaarden van E zijn, zodat voor alle j de $D_{j\lambda}$ omkeerbaar zijn. Mits ook D_l omkeerbaar is, is V_j uitdrukbaar in een nulvector $y_{\lambda i} \neq 0$ van $C_{i\lambda}$:

$$\boxed{V_{j\lambda} = -D_{j\lambda}^{-1} y_{\lambda i} \quad v_{i\lambda} = D_l^{-1} y_{\lambda i} \Rightarrow C_{i\lambda} y_{\lambda i} = 0 \quad C_{i\lambda} = D_l^{-1} + \Sigma'_j D_{j\lambda}^{-1}}$$

De waarden van λ waarvoor $C_{i\lambda}$ singulier wordt (dwz een nulruimte heeft) zijn de eigenwaarden van F.

Bestaan de omgekeerden van E, D_j en D_l, symbolisch F genaamd? Neem vooralsnog aan dat λ zodanig is dat alle omgekeerden bestaan. We bepalen de omgekeerde van 'een' matrix F met hoofdstruktuur als boven geschetst, dus met 'submatrices' (eventueel scalars) genaamd D_l en D_j. Eerst vegen we de rijen tot een diagonale hoofdstructuur; dat gaat dus zonder verandering van de determinant. Daarna diagonalizeren we de hoofddiagonaal.

Start met $FF^{-1} = 1$.

$$\begin{pmatrix} D_1{+}D_l & D_l & D_l & D_l & \cdots \\ D_l & D_2 + D_l & D_l & D_l & \cdots \\ D_l & D_l & D_3 + D_l & D_l & \cdots \\ D_l & D_l & D_l & D_4 + D_l & \cdots \\ \cdots \end{pmatrix} F^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \cdots \end{pmatrix}$$

Verminder iedere rij met de eerste rij:

$$\begin{pmatrix} D_1{+}D_l & D_l & D_l & D_l & \cdots \\ -D_1 & D_2 & 0 & 0 & \cdots \\ -D_1 & 0 & D_3 & 0 & \cdots \\ -D_1 & 0 & 0 & D_4 & \cdots \\ \cdots \end{pmatrix} F^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ -1 & 1 & 0 & 0 & \cdots \\ -1 & 0 & 1 & 0 & \cdots \\ -1 & 0 & 0 & 1 & \cdots \\ \cdots \end{pmatrix}$$

Verminder de eerste rij met $D_l D_j^{-1}$ maal rij-j voor alle $j > 1$; met
$C = D_l^{-1}+D_1^{-1}+D_2^{-1}+D_3^{-1}+D_4^{-1}+\cdots$

$$\begin{pmatrix} D_l C D_1 & 0 & 0 & 0 & \cdots \\ -D_1 & D_2 & 0 & 0 & \cdots \\ -D_1 & 0 & D_3 & 0 & \cdots \\ -D_1 & 0 & 0 & D_4 & \cdots \\ \cdots \end{pmatrix} F^{-1} = \begin{pmatrix} D_l C - D_l D_1^{-1} & -D_l D_2^{-1} & -D_l D_3^{-1} & -D_l D_4^{-1} & \cdots \\ -1 & 1 & 0 & 0 & \cdots \\ -1 & 0 & 1 & 0 & \cdots \\ -1 & 0 & 0 & 1 & \cdots \\ \cdots \end{pmatrix}$$

Vermeerder alle (behalve eerste) rijen met $C^{-1}D_l^{-1}$ maal de eerste rij:

$$\begin{pmatrix} D_l C D_1 & 0 & 0 & \cdots \\ 0 & D_2 & 0 & \cdots \\ 0 & 0 & D_3 & \cdots \\ \cdots \end{pmatrix} F^{-1} = \begin{pmatrix} D_l C - D_l D_1^{-1} & -D_l D_2^{-1} & -D_l D_3^{-1} \cdots \\ -C^{-1}D_1^{-1} & 1-C^{-1}D_2^{-1} & -C^{-1}D_3^{-1} \cdots \\ -C^{-1}D_1^{-1} & -C^{-1}D_2^{-1} & 1-C^{-1}D_3^{-1} \cdots \\ \cdots \end{pmatrix}$$

Iedere j-cel van de hoofddiagonaal bevat bijpassende eigenwaarden. De determinant $|F| = |D_l||C||D_1||D_2||D_3||D_4|\cdots$. Diagonaliseer door iedere rij te vermenigvuigen met de inverse diagonaalcel: $D_1^{-1}C^{-1}D_l^{-1}$, D_2^{-1}, D_3^{-1}, enz.

$$\begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \cdots \end{pmatrix} F^{-1} = \begin{pmatrix} D_1^{-1}-D_1^{-1}C^{-1}D_1^{-1} & -D_1^{-1}C^{-1}D_2^{-1} & -D_1^{-1}C^{-1}D_3^{-1}\cdots \\ -D_2^{-1}C^{-1}D_1^{-1} & D_2^{-1}-D_2^{-1}C^{-1}D_2^{-1} & -D_2^{-1}C^{-1}D_3^{-1}\cdots \\ -D_3^{-1}C^{-1}D_1^{-1} & -D_3^{-1}C^{-1}D_2^{-1} & D_3^{-1}-D_3^{-1}C^{-1}D_3^{-1}\cdots \\ \cdots \end{pmatrix}$$

De laatste matrix rechts is de gezochte omgekeerde F^{-1} van F! Merk op, dat D_l alleen als omgekeerde in C voorkomt! De matrix en omgekeerde blijken te schrijven als:

$$\boxed{F = \delta_{jj'}D_j+D_l \quad F^{-1} = \delta_{jj'}D_j^{-1}-D_j^{-1}C^{-1}D_{j'}^{-1} \qquad C = \Sigma_j D_j^{-1}}$$

Merk op dat, als we rijen èn kolommen van F^{-1} vergroten met D_j, de matrix $D_jF^{-1}D_{j'} = \delta_{jj'}D_j - C^{-1}$ weer van het zelfde type is, dus omkeerbaar, met D_l vervangen door $-C^{-1}$. Bijgevolg is bij herhaling hiervan $(D_jF^{-1}D_{j'})^{-1} = \delta_{jj'}D_j^{-1}-D_j^{-1}((-C^{-1})^{-1}+\Sigma_j' D_j^{-1})^{-1}D_{j'}^{-1} = \delta_{jj'}D_j^{-1} - D_j^{-1}(-D_l^{-1})^{-1}D_{j'}^{-1}$; we zijn inderdaad terug bij $F = \delta_{jj'}D_j + D_l$!

De omkeereigenschappen van de 'elementen' in de F-matrices zijn als volgt:

$\mathbf{D_j}$ van 'F zonder λ' inclusief $j = l$:

- $c = 1/d_1+\cdots+1/d_{k-1}+1/d_k = \sum_i{}' P^{\mathrm{h}}_{ij}+P^{\mathrm{h}}_{kj} = P_j \neq 0$;
- determinant $|D_j| = d_1\cdots d_{k-1}c = (1/P^{\mathrm{h}}_{1j}\cdots 1/P^{\mathrm{h}}_{kj})P_j \neq 0$; D_j^{-1} bestaat;
- omgekeerde $\boxed{{D_j^{-1}}_{ii'} = \delta_{i,i'}P^{\mathrm{h}}_{ij} - P^{\mathrm{h}}_{ij}P^{\mathrm{h}}_{i'j}/P_j}$;
- som omgekeerden $\boxed{C_{0ii'} = \delta_{i,i'}P_i - \Sigma_j P^{\mathrm{h}}_{ij}P^{\mathrm{h}}_{i'j}/P_j}$

$\mathbf{D_{j\lambda}}$ van '$F - 1/\lambda$' of E; Gebruik de diagonaalterm $d_i = 1/P^{\mathrm{h}}_{ij}-1/\lambda = -1/A_{ij\lambda}$

- $c = 1/d_1+\cdots+1/d_{k-1}+1/d_k = \sum_i{}'(-A_{ij\lambda})+P^{\mathrm{h}}_{kj} = \boxed{(P^{\mathrm{h}}_{kj} - A_{j\lambda}) = B_{j\lambda}}$;
- determinant $|D_{j\lambda}| = d_1\cdots d_{k-1}c = (1/P^{\mathrm{h}}_{1j}-1/\lambda)\cdots(1/P^{\mathrm{h}}_{k-1j}-1/\lambda)B_{j\lambda}$;
- omgekeerde bestaat als $A_{j\lambda} \neq P^{\mathrm{h}}_{kj}$: $\boxed{{(D_{j\lambda})^{-1}}_{ii'} = -\delta_{i,i'}A_{ij\lambda} - A_{ij\lambda}B_{j\lambda}^{-1}A_{i'j\lambda}}$;

– Als $B_{j\lambda} \neq 0$, of $A_{j\lambda} \neq P^{\mathrm{h}}_{kj}$, bestaat $D^{-1}_{j\lambda}$. Als dat geldt voor alle $j = 1 \cdots l$ dan bestaat $C_{i\lambda}$:

$$\boxed{C_{i\lambda ii'} = (\delta_{i,i'} B_{i\lambda} - \Sigma_j{}' A_{ij\lambda} B^{-1}_{j\lambda} A_{i'j\lambda}) - P^{\mathrm{h}}_{il} P^{\mathrm{h}}_{i'l}/P_l)}$$

De eigenwaarden van F worden bepaald door het verdwijnen van de determinant van de $E = F - 1/\lambda$-matrix, die van het F type is. Uit bovenstaande afleiding van de omgekeerde blijkt

$$\boxed{\det(E) = (\Pi'_j |D_{j\lambda}|)|D_l|\det(C_{i\lambda})} \quad \boxed{\det(C_{i\lambda}) = \det(E)/(\Pi'_j |D_{j\lambda}|)|D_l|}$$

zodat de eigenwaarden kunnen worden bepaald uit het verdwijnen van de determinant van $C_{i\lambda} = D^{-1}_l + \Sigma'_j D^{-1}_{j\lambda}$. Uit de definitie van C volgt dat die determinant oneindig wordt als een $D_{j\lambda}$ singulier is, dus als de eigenwaarde ook een eigenwaarde van D_j is.

De oplossing $y_{i\lambda}$ is nulvector van de symmetrische matrix $C_{i\lambda}$. Omgekeerd leidt een nulvector y_i van C_i met $v_{i\lambda} = D^{-1}_l y_{i\lambda}$ en $v_{ij\lambda} = -D^{-1}_{j\lambda} y_{i\lambda}$ tot de deelsom variabelen:

$$\begin{aligned}
v_{i\lambda} &= D_l{}^{-1} y_i = P^{\mathrm{h}}_{il}(y_{i\lambda} - (1/P_l)(\sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda})) \\
v_\lambda &= \sum_i{}' P^{\mathrm{h}}_{il} y_{\lambda i}(1 - \sum_i{}' P^{\mathrm{h}}_{il}/P_l) = (\sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda}) P^{\mathrm{h}}_{kl}/P_l = P^{\mathrm{h}}_{kl} x_\lambda \\
&\to \boxed{x_\lambda = (1/P_l)\sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda}} \\
x_{i\lambda} &= v_{i\lambda}/P^{\mathrm{h}}_{il} = y_{i\lambda} - (1/P_l)\sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda} \to \boxed{x_{i\lambda} = y_{i\lambda} - x_\lambda} \\
v_{ij\lambda} &= -D^{-1}_{j\lambda} y = \Sigma_{i'}{}' A_{ij}(\delta_{ii'} + B^{-1}_j A_{i'j}) y_{i'\lambda} = A_{ij}(y_{i\lambda} + B^{-1}_j \Sigma_{i'}{}' A_{i'j} y_{i'\lambda}) \\
v_{j\lambda} &= \Sigma_i{}' A_{ij}(y_{i\lambda} + B^{-1}_j \Sigma_{i'}{}' A_{i'j} y_{i'\lambda}) = (1 + A_j B^{-1}_j)\Sigma_{i'}{}' A_{ij} y_{i'\lambda} = P^{\mathrm{h}}_{kj} B^{-1}_j \Sigma_i{}' A_{ij} y_{i\lambda} \\
x_{j\lambda} &= v_{j\lambda}/P^{\mathrm{h}}_{kj} \to \boxed{x_{j\lambda} = B^{-1}_j \Sigma_i{}' A_{ij} y_{i\lambda}} \\
&\to \boxed{v_{ij\lambda} = A_{ij}(y_{i\lambda} + x_{j\lambda}) = A_{ij}(x_{i\lambda} + x_{j\lambda} + x_\lambda)}
\end{aligned}$$

Hiermee is aangetoond dat alle oplossingen van de eigenvectoren $v_{ij\lambda}$ van F uit de nulvectoren y_i van C_i volgen. Het vinden van de waarden waarvoor $C_{i\lambda}$ singulier wordt, en de bijbehorende nulvectoren, is een alternatief voor het bepalen van eigenwaarden en eigenvectoren van F.

Naast deze oplossing is evengoed de duale oplossing, door verwisseling van i en j (dus ook k en l). De duale oplossing $y_{j\lambda}$ in de j-ruimte is nulvector van de duale matrix C_j in de j-ruimte:

$$\sum_{j'}{}'((\delta_{j,j'} B_j - \sum_i{}' A_{ij} B^{-1}_i A_{ij'}) - P^{\mathrm{h}}_{kj} P^{\mathrm{h}}_{kj'}/P_k) y_{\lambda j'} = 0 \quad V_j = -D^{-1}_{j\lambda} y_{j\lambda}$$

De duale vergelijking zal dezelfde oplossing geven als de gewone vergelijking.
2017: Het verband tussen de oplossing $y_{i\lambda}$ in de i-ruimte en de duale oplossing $y_{j\lambda}$ in de j ruimte is:

$$\begin{aligned}
y_{j\lambda} &= x_{j\lambda} + x_\lambda = \Sigma_i{}' S_{ji} y_{i\lambda} \quad S_{ji} = B^{-1}_j A^t_{ij} + P^{-1}_l P^{\mathrm{h}}_{il} \\
y_{i\lambda} &= x_{i\lambda} + x_\lambda = \Sigma_j{}' T_{ij} y_{j\lambda} \quad T_{ij} = B^{-1}_i A_{ij} + P^{-1}_k P^{\mathrm{h}}_{kj} \\
&\Rightarrow \Sigma_i{}' S_{ji} T_{ij'} = \delta_{jj'} \quad \Sigma_j{}' T_{ij} S_{ji'} = \delta_{ii'}
\end{aligned}$$

We zien dat S_{ij} en T_{ij} elkaars omgekeerde zijn, een direct gevolg van de dualiteit.

3.3.3 'Oplossing' met gemengde variabelen $x_{i\lambda}$ en $y_{j\lambda}$

2015: Op grond van het vorige subhoofdstuk geven we de oplossing van de eigenvectorvergelijking voor v opnieuw, maar nu met gemengde variabelen. Veronderstel weer dat de eigenwaarden λ bekend zijn, bijvoorbeeld uit determinant onderzoek. Naast $x_{i\lambda}$, $x_{j\lambda}$ en x_λ zijn nieuwe variabelen $y_{i\lambda}$ en $y_{j\lambda}$:

$$x_{i\lambda} = v_{i\lambda}/P^{\mathrm{h}}_{il} \quad x_{j\lambda} = v_{j\lambda}/P^{\mathrm{h}}_{kj} \quad x_\lambda = v_\lambda/P^{\mathrm{h}}_{kl} \quad y_{i\lambda} = x_{i\lambda} + x_\lambda \qquad y_{j\lambda} = x_{j\lambda} + x_\lambda$$

$$A_{\lambda,ij} = \frac{1}{1/\lambda - 1/P^{\mathrm{h}}_{ij}} \quad v_{ij\lambda} = A_{ij}(x_{i\lambda} + x_{j\lambda} + x_\lambda) = A_{ij}(y_{i\lambda} + x_{j\lambda}) = A_{ij}(y_{j\lambda} + x_{i\lambda})$$

Met deelsommaties verkrijgen we $k{+}l{-}1$ vergelijkingen voor evenzoveel onbekenden:

$$v_{i\lambda} = P^{\mathrm{h}}_{il}x_{i\lambda} = P^{\mathrm{h}}_{il}(y_{i\lambda}{-}x_\lambda) = A_i y_{i\lambda} + \Sigma'_j A_{ij}x_{j\lambda} = \Sigma'_j A_{ij}y_{j\lambda} + A_i x_{i\lambda}$$

$$v_{j\lambda} = P^{\mathrm{h}}_{kj}x_{j\lambda} = \Sigma'_i A_{ij}y_{i\lambda} + A_j x_{j\lambda} = A_j y_{j\lambda} + \Sigma'_i A_{ij}x_{i\lambda} = P^{\mathrm{h}}_{kj}(y_{j\lambda}{-}x_\lambda)$$

$$v_\lambda = P^{\mathrm{h}}_{kl}x_\lambda = \Sigma'_i P^{\mathrm{h}}_{il}x_{i\lambda} = \Sigma'_i A_i y_{i\lambda} + \Sigma'_j A_j x_{j\lambda} = \Sigma'_j A_j y_{j\lambda} + \Sigma'_i A_i x_{i\lambda} = \Sigma'_j P^{\mathrm{h}}_{kj}x_{j\lambda}$$

$$\Rightarrow \Sigma'_i P^{\mathrm{h}}_{il}y_{i\lambda} = \Sigma'_i P^{\mathrm{h}}_{il}(x_{i\lambda}{+}x_\lambda) = \Sigma'_i P^{\mathrm{h}}_{il}x_{i\lambda}{+}(P_l - P^{\mathrm{h}}_{kl})x_\lambda = P_l x_\lambda$$

$$\Rightarrow \Sigma'_j P^{\mathrm{h}}_{kj}y_{j\lambda} = \Sigma'_j P^{\mathrm{h}}_{kj}(x_{j\lambda}{+}x_\lambda) = P^{\mathrm{h}}_{kl}x_\lambda + (P_k - P^{\mathrm{h}}_{kl})x_\lambda = P_k x_\lambda$$

Terzijde merken we op, dat er een rechtstreeks, en omkeerbaar, verband is tussen $v_{i\lambda}$ en $y_{i\lambda}$, en duaal:

$$v_{i\lambda} = P^{\mathrm{h}}_{il}(y_{i\lambda}{-}x_\lambda) = P^{\mathrm{h}}_{il}(y_{i\lambda}{-}\Sigma'_{i'}P^{\mathrm{h}}_{i'l}y_{i'\lambda}/P_l) = \Sigma'_{i'}(\delta_{ii'}P^{\mathrm{h}}_{il}{-}P^{\mathrm{h}}_{il}P^{\mathrm{h}}_{i'l}/P_l)y_{i'\lambda}$$

$$v_{i\lambda} = D_l^{-1}y_{i\lambda} \qquad v_{j\lambda} = D_k^{-1}y_{j\lambda}$$

Bovenstaande vergelijkingen voor $x_{i\lambda}$, $y_{j\lambda}$ enz. vereenvoudigen, met B_i en B_j, tot:

$$(P^{\mathrm{h}}_{il}{-}A_i)x_{i\lambda} = \Sigma'_j A_{ij}y_{j\lambda} \to B_i x_{i\lambda} = \Sigma'_j A_{ij}y_{j\lambda} \quad B_i = P^{\mathrm{h}}_{il}{-}A_i$$

$$(P^{\mathrm{h}}_{il} - A_i)y_{i\lambda}{-}P^{\mathrm{h}}_{il}x_\lambda = \Sigma'_j A_{ij}x_{j\lambda} \to B_i y_{i\lambda} - P^{\mathrm{h}}_{il}x_\lambda = \Sigma'_j A_{ij}x_{j\lambda}$$

$$(P^{\mathrm{h}}_{kj}{-}A_j)x_{j\lambda} = \Sigma'_i A_{ij}y_{i\lambda} \to B_j x_{j\lambda} = \Sigma'_i A_{ij}y_{i\lambda} \quad B_j = P^{\mathrm{h}}_{kj}{-}A_j$$

$$(P^{\mathrm{h}}_{kj}{-}A_j)y_{j\lambda} - P^{\mathrm{h}}_{kj}x_\lambda = \Sigma'_i A_{ij}x_{i\lambda} \to B_j y_{j\lambda} - P^{\mathrm{h}}_{kj}x_\lambda = \Sigma'_i A_{ij}x_{i\lambda}$$

en na uitdrukken van x_λ in $y_{i\lambda}$ of $y_{j\lambda}$:

$$B_i y_{i\lambda} - P^{\mathrm{h}}_{il}\Sigma'_i P^{\mathrm{h}}_{il}y_{i\lambda}/P_l = \Sigma'_j A_{ij}x_{j\lambda}$$

$$B_j y_{j\lambda} - P^{\mathrm{h}}_{kj}\Sigma'_j P^{\mathrm{h}}_{kj}y_{j\lambda}/P_k = \Sigma'_i A_{ij}x_{i\lambda}$$

Als $B_i \neq 0$ kan $x_{i\lambda}$ worden uitgedrukt in $y_{j\lambda}$; als $B_j \neq 0$ kan $x_{j\lambda}$ worden uitgedrukt in $y_{i\lambda}$. Daaruit volgen omgekeerd gekoppelde vergelijkingen voor $y_{i\lambda}$ en $x_{j\lambda}$ (duaal: $y_{j\lambda}$ en $x_{i\lambda}$). Dan blijken $y_{i\lambda}$ en $y_{j\lambda}$ nuleigenvectoren van duale symmetrische matrices $C_{\lambda i}$ en $C_{j\lambda}$:

$$x_{j\lambda} = B_j^{-1}\Sigma'_i A_{ij}y_{i\lambda} \quad C_{\lambda ii'} = \delta_{ii'}B_i - P^{\mathrm{h}}_{il}P^{\mathrm{h}}_{i'l}/P_l - \Sigma'_j A_{ij}B_j^{-1}A_{i'j} \quad \Sigma'_{i'}C_{\lambda ii'}y_{i'\lambda} = 0$$

$$x_{i\lambda} = B_i^{-1}\Sigma'_j A_{ij}y_{j\lambda} \quad C_{\lambda jj'} = \delta_{jj'}B_j - P^{\mathrm{h}}_{kj}P^{\mathrm{h}}_{kj'}/P_k - \Sigma'_i A_{ij}B_i^{-1}A_{ij'} \quad \Sigma'_{j'}C_{\lambda jj'}y_{j'\lambda} = 0$$

$$v_{ij\lambda} = A_{ij}y_{i\lambda} + A_{ij}B_j^{-1}\Sigma_{i'}A_{i'j}y_{i'\lambda} = A_{ij}y_{j\lambda} + A_{ij}B_i^{-1}\Sigma_{j'}A_{ij'}y_{j'\lambda}$$

Merk op, dat een van beide, $y_{i\lambda}$ of $y_{j\lambda}$, voldoende is voor het oplossen van $v_{ij\lambda}$: ze zijn duaal gelijkwaardig. $y_{i\lambda}$, een nulvector van $C_{\lambda i}$, is gekoppeld aan $x_{j\lambda}$ en x_λ. Deze drie zijn gekoppeld aan x_λ, $x_{i\lambda}$ en $y_{j\lambda}$, die een nulvector van $C_{j\lambda}$ is. Dus zijn de nulruimten van $C_{i\lambda}$ en $C_{j\lambda}$, bepaald door een eigenwaarde λ en eigenvector v_λ, gekoppeld en even groot (van dezelfde dimensie). Naast die gekoppelde nulruimte zal $C_{j\lambda}$ een nulruimte hebben van dimensie $l - k$, als de dimensie l van de j-tuimte groter is dan de dimensie k van de i-ruimte.

[2017, 2019] Merk verder op, dat in C_i de symmetrische matrix $\overline{B_i}$ voorkomt:

$$C_{i\lambda} = (\delta_i P^{\mathrm{h}}_{il} - P^{\mathrm{h}}_{il}{P^{\mathrm{h}}_{il}}^t/P_l) + (A + AB_j^{-1}A^t)$$

$$\overline{B_i} = B_i - AB_j^{-1}A^t \quad B_i = \delta_i P^{\mathrm{h}}_{il} - Au_j$$

$$\to \overline{B_i} = C_i + P^{\mathrm{h}}_{il}{P^{\mathrm{h}}_{il}}^t/P_l$$

Omdat $y_{i\lambda}$ een nulvector is van C_i geldt, als $\overline{B}$ omkeerbaar is:

$$\overline{B_i} y_{i\lambda} = P^{\mathrm{h}}_{il} {P^{\mathrm{h}}_{il}}^t y_{i\lambda}/P_l = P^{\mathrm{h}}_{il} x$$

$$y_{i\lambda} = \overline{B_i}^{-1} P^{\mathrm{h}}_{il} x \qquad y_{j\lambda} = \overline{B_j}^{-1} P^{\mathrm{h}}_{kj} x$$

$$x_{i\lambda} = (\overline{B_i}^{-1} P^{\mathrm{h}}_{il} - u_i)x \qquad x_{j\lambda} = (\overline{B_j}^{-1} P^{\mathrm{h}}_{kj} - u_j)x$$

en duaal. De consistentievergelijking wordt:

$$\boxed{{P^{\mathrm{h}}_{il}}^t \overline{B_i}^{-1} P^{\mathrm{h}}_{il} = P_l} \qquad \boxed{{P^{\mathrm{h}}_{kj}}^t \overline{B_j}^{-1} P^{\mathrm{h}}_{kj} = P_k}$$

Het verband tussen $\overline{B_{i\lambda}}$, $C_{i\lambda}$ en P^{h}_{il} laat zien dat ook C_i als Schur-complement kan worden gezien, maar dan in de matrix van de vergelijkingen:

$$\begin{pmatrix} \overline{B_i} & -P^{\mathrm{h}}_{il} \\ -{P^{\mathrm{h}}_{il}}^t & P_l \end{pmatrix} \begin{pmatrix} y_{i\lambda} \\ x_\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Vergelijk dat met de Schur-complementen in vierkante matrix AB/CD:

$$\overline{A} = A - BD^{-1}C \qquad \overline{D} = D - CA^{-1}B$$

De transformatie: $A \to \overline{B_i}$, $B \to -P^{\mathrm{h}}_{il}$, $C \to -{P^{\mathrm{h}}_{il}}^t$, $D \to P_l$, maakt in die matrix C_i het complement van $\overline{B_i}$:

$$C_i = \mathrm{compl}(\overline{B_i}) = \overline{B_i} - P^{\mathrm{h}}_{il} P_l^{-1} {P^{\mathrm{h}}_{il}}^t \qquad \mathrm{compl}(P_l) = P_l - {P^{\mathrm{h}}_{il}}^t \overline{B_i}^{-1} P^{\mathrm{h}}_{il},$$

en de i-consistentievergelijking het complement van P_l. Daardoor is het singulier zijn van C_i gekoppeld aan het verdwijnen (nul zijn) van $\mathrm{compl}(P_l) = P_l - {P^{\mathrm{h}}_{il}}^t \overline{B_i}^{-1} P^{\mathrm{h}}_{il}$. Een algemene eigenschap van complementomkering [6] geeft het indirecte verband tussen die twee:

$$(1 - CA^{-1}BD^{-1})^{-1} = (1 + C\overline{A}^{-1}BD^{-1})$$

$$\to (1 - {P^{\mathrm{h}}_{il}}^t \overline{B_i}^{-1} P^{\mathrm{h}}_{il}/P_l)^{-1} = (1 + {P^{\mathrm{h}}_{il}}^t C_i^{-1} P^{\mathrm{h}}_{il}/P_l)$$

3.3.4 Functieonderzoek $A_{ij\lambda}$ en $B_{i\lambda}$

Voor het bepalen van de eigenwaarden kunnen we de vergelijking $|C_{i\lambda}| = 0$ voor λ oplossen. Dat is een complexe functie van sommen van producten van factoren als $A_{ij\lambda}$, $B_{i\lambda}$ en $B_{j\lambda}$, waarin λ. Daarbij zijn te onderscheiden

$$A_{ij} = \frac{1}{1/\lambda - 1/P^{\mathrm{h}}_{ij}} = \frac{{P^{\mathrm{h}}_{ij}}^2}{P^{\mathrm{h}}_{ij} - \lambda} - P^{\mathrm{h}}_{ij}$$

$$A_i = \sum_j{}' \frac{{P^{\mathrm{h}}_{ij}}^2}{P^{\mathrm{h}}_{ij} - \lambda} - (P_i - P^{\mathrm{h}}_{il}) \qquad A_j = \sum_i{}' \frac{{P^{\mathrm{h}}_{ij}}^2}{P^{\mathrm{h}}_{ij} - \lambda} - (P_j - P^{\mathrm{h}}_{kj})$$

$$B_i = P_i - \sum_j{}' \frac{{P^{\mathrm{h}}_{ij}}^2}{P^{\mathrm{h}}_{ij} - \lambda} \qquad B_j = P_j - \sum_i{}' \frac{{P^{\mathrm{h}}_{ij}}^2}{P^{\mathrm{h}}_{ij} - \lambda}$$

$$A = \sum_{ij}{}' \frac{{P^{\mathrm{h}}_{ij}}^2}{P^{\mathrm{h}}_{ij} - \lambda} - (1 - P_k - P_l + P^{\mathrm{h}}_{kl})$$

$$B = \sum_{ij}{}' \frac{{P^{\mathrm{h}}_{ij}}^2}{P^{\mathrm{h}}_{ij} - \lambda} + (1 - P_k - P_l + 2P^{\mathrm{h}}_{kl})$$

Dus is A_{ij} een stijgende functie met singulariteit: verticale asymptoot $\lambda = P^{\mathrm{h}}_{ij}$ bij $A_{ij} = \pm\infty$, horizontale asyptoot $A_{ij} = -P^{\mathrm{h}}_{ij}$ bij $\lambda = \mp\infty$ en nuldoorgang $A_{ij} = 0$ bij $\lambda = 0$. Verder

[6] zie de appendix Eigenschappen van Schur-complementen.

zijn A_i respectievelijk A_j stijgende functies, want sommen van (A_{ij}) stijgende functies, met horizontale asyptoot $A_i = -(P_i - P^{\mathrm{h}}_{il})$ resp. $A_j = -(P_j - P^{\mathrm{h}}_{kj})$, en maximaal $k-1$ resp. $l-1$ $A_{i,j}$ singulariteiten bij $\lambda = P^{\mathrm{h}}_{i,j1}$ resp. $\lambda = P^{\mathrm{h}}_{i1,j}$ —met evenzovele intervallen met nuldoorgangen.
Analoog zijn $B_i = P^{\mathrm{h}}_{il} - A_i$ en $B_j = P^{\mathrm{h}}_{kj} - A_j$ dalende functies met singulariteiten $B_i = \mp\infty$ resp. $B_j = \mp\infty$, bij de $l-1$ $\lambda = P^{\mathrm{h}}_{ij}$ resp. de $k-1$ $\lambda = P^{\mathrm{h}}_{ij}$; de horizontale asyptoot is $B_i = P_i$ resp. $B_j = P_j$. Op ieder interval, tussen opvolgende singulariteiten $P^{\mathrm{h}}_{i\cdot}$, ligt een nuldoorgang van B_i, resp. voor B_j, alle dus tussen 0 en 1. Omdat bij $\lambda = 0$ de $B_i = P^{\mathrm{h}}_{il} > 0$, is ook de nuldoorgang van het eerste interval positief. De omgekeerde $1/B_i$ en $1/B_j$ zijn dan weer stijgende functies, met nuldoorgangen $\lambda = P^{\mathrm{h}}_{ij}$ daar waar A_i resp. A_j een singulariteit heeft. De singulariteiten van $1/B_i$ en $1/B_j$ zitten in de intervallen tussen de nuldoorgangen $\lambda = P^{\mathrm{h}}_{ij}$; de horizontale asyptoot is $1/B_i = 1/P_i$ en $1/B_j = 1/P_j$.

Bij producten van deze functies moeten we rekening houden met het teken van de functie: de negatie van stijgend is dalend en andersom. Om te beginnen zien we dat de producten A_{ij}/B_j en $A_{ij}/B_i = A_{ij}/(P^{\mathrm{h}}_{il} - A_i)$ voorkomen in de sommen $\sum_j{}' A_{ij}/B_j$ en $\sum_i{}' A_{ij}/B_i$. De functie A_j is een i'-som van $A_{i',j}$ termen, waaronder A_{ij}. Als gevolg daarvan is de singulariteit $\lambda = P^{\mathrm{h}}_{ij}$ van A_{ij} geen singulariteit van A_{ij}/B_j:

$$A_{ij}/B_j = A_{ij}/(P^{\mathrm{h}}_{kj} - \Sigma''_{i'} A_{i'j} - A_{ij}) \to -1 \text{ als } A_{ij} \to \pm\infty$$

De singulariteiten van A/B_j zijn dus de nuldoorgangen van de B_j, bepaald door de singulariteiten van $1/B_j$ zoals eerder beschreven.

Vervolgens de productsom $\sum_j{}' A_{ij} B_j^{-1} A_{i'j}$. Bij een niet-diagonaal element, $i \neq i'$, kunnen we het product van A factoren omzetten in het verschil:

$$A_{ij}A_{i'j} = (1/P^{\mathrm{h}}_{i'j} - 1/P^{\mathrm{h}}_{ij})(A_{ij} - A_{i'j})$$

In de som A_j komen beide termen A_{ij} en $A_{i'j}$ voor. zodat $A_{ij}B_j^{-1}A_{i'j}$ dezelfde singulariteiten heeft als AB_j^{-1}: de nuldoorgangen van de B_j. Bij een diagonaal element, $i = i'$, daarentegen is de singulariteit van $A^2_{i,j}$ in de teller sterker dan de singulariteit van $A_{i,j}$ in de noemer, dus in B_j, zodat naast de singulariteiten van B_j ook die van $A_{i,j}$ een rol speelt.

$$A^2_{ij}/B_j = A^2_{ij}/(P^{\mathrm{h}}_{kj} - \Sigma''_{i'} A_{i'j} - A_{ij}) \to \mp\infty \text{ als } A_{ij} \to \pm\infty$$

Tenslotte de termen in C_i die van λ afhangen: $B_i - AB_j^{-1}A^t$. De niet-diagonale elementen zijn $-AB_j^{-1}A^t$, zojuist behandeld. Bij de niet diagonaal elementen komt de singulariteit van A_{ij} voor in A, B_j en A^t zoals geschetst. De leidende term van A_{ij} rond $\lambda = P^{\mathrm{h}}_{ij}$:

$$-A_{ij}A_{i'j}/((P^{\mathrm{h}}_{kj} - \Sigma''_{i'} A_{i'j}) - A_{ij}) \to -A_{i'j}$$

Bij de diagonaalelementen komt de singulariteit van A_{ij} voor in B_i en in $AB_j^{-1}A^t$:

$$B_i - A^2_{ij}/B_j = (P^{\mathrm{h}}_{il} - \Sigma''_{j'} A_{ij} - A_{ij}) - (\Sigma''_{j'} A^2_{ij'}/B_{j'} + A^2_{ij}/((P^{\mathrm{h}}_{kj} - \Sigma''_{i'} A_{i'j}) - A_{ij})$$

Noem $(P^{\mathrm{h}}_{il} - \Sigma''_{j'} A_{ij'}) = b_i$, $(P^{\mathrm{h}}_{kj} - \Sigma''_{i'} A_{i'j}) = b_j$, $\Sigma''_{j'} A^2_{ij'}/B_{j'} = aba$ en $A_{ij} = x$ met $x \to \pm\infty$. Dan wordt de leidende term bij de singulariteit van A_{ij} rond $\lambda = P^{\mathrm{h}}_{ij}$:

$$\begin{aligned}
&b_i - x - aba - x^2/(b_j - x) = [(b_i - aba - x)(b_j - x) - x^2]/(b_j - x)\\
&= [(b_i - aba)b_j - (b_i - aba + b_j)x]/(b_j - x) \to (b_i - aba + b_j) + b_j^2/x\\
\to&(P^{\mathrm{h}}_{il} - \Sigma''_{j'} A_{ij'}) + (P^{\mathrm{h}}_{kj} - \Sigma''_{i'} A_{i'j}) - \Sigma''_{j'} A^2_{ij'}/B_{j'} + \cdots
\end{aligned}$$

3.4 Kansen bij de eigenvectoren

Omdat de eigenvectoren geen bijdrage hebben bij $i = k$ of $j = l$ nemen we $v_{kj\lambda} = v_{il\lambda} = v_{kl\lambda} = 0$, dus ook $x_{k\lambda} = x_{l\lambda} = 0$. De kans P_λ op 'oorzaaken/of gevolg' met eigenwaarde λ en eigenvector $v_{ij\lambda}$ is bij definitie $P(v_{ij\lambda}) = \sum_{ij} P_{ij} v_{ij\lambda} = \sum_{ij}{}' P_{ij} v_{ij\lambda}$. Kenmerkend voor de eigenvector $v_{ij\lambda}$ is de factor A: $v_{ij\lambda} = A_{ij}(y_{i\lambda}+x_{j\lambda})$ met $A_{ij} = 1/(1/\lambda - 1/P^{\mathrm{h}}_{ij})$. Voor A geldt de 'interessante eigenschap':

$$1/\lambda = 1/A_{ij} + 1/P^{\mathrm{h}}_{ij} = (P^{\mathrm{h}}_{ij}+A_{ij})/A_{ij}P^{\mathrm{h}}_{ij} \quad P^{\mathrm{h}}_{ij}A_{ij} = \lambda(A_{ij}+P^{\mathrm{h}}_{ij})$$

$$P^{\mathrm{h}}_{ij}v_{ij\lambda} = \lambda(v_{ij\lambda}+P^{\mathrm{h}}_{ij}(y_{i\lambda}+x_{j\lambda}))$$

Daarmee kunnen kans op en covariantie van de eigenvectoren gedeeltelijk worden bepaald.

$$\begin{aligned}
P_\lambda &= \sum_{ij}{}' P_{ij}v_{ij\lambda} = \sum_{ij}{}'(P^{\mathrm{h}}_{ij}+e_{ij})v_{ij\lambda} = \sum_{ij}{}' P^{\mathrm{h}}_{ij}v_{ij\lambda} + \sum_{ij}{}' e_{ij}v_{ij\lambda} = P^{\mathrm{h}}_\lambda + e_\lambda \\
P^{\mathrm{h}}_\lambda &= \sum_{ij}{}' P^{\mathrm{h}}_{ij}A_{ij}(y_{i\lambda}+x_{j\lambda}) = \sum_{ij}{}' \lambda(A_{ij}+P^{\mathrm{h}}_{ij})(y_{i\lambda}+x_{j\lambda}) \\
&= \lambda\sum_{ij}{}'(v_{ij\lambda}+P^{\mathrm{h}}_{ij}(y_{i\lambda}+x_{j\lambda})) = \lambda(v_\lambda + P^{\mathrm{h}}(y_{i\lambda}+x_{j\lambda}))
\end{aligned}$$

$$P^{\mathrm{h}}(y_{i\lambda}) = \sum_i{}'(\sum_j{}' P^{\mathrm{h}}_{ij})y_{i\lambda} \quad P^{\mathrm{h}}(x_{j\lambda}) = \sum_j{}'(\sum_i{}' P^{\mathrm{h}}_{ij})x_{j\lambda} \quad P^{\mathrm{h}}(x) = (\sum_i{}'\sum_j{}' P^{\mathrm{h}}_{ij})x$$

Merk op, dat $P^{\mathrm{h}}(y_{i\lambda})$ en $P^{\mathrm{h}}(x_{j\lambda})$ niet (meer) van index i of j afhangen. De volgende deelsommen van P^{h}_{ij} zijn te herkennen:

$$\sum_i{}' P^{\mathrm{h}}_{ij} = (P_j - P^{\mathrm{h}}_{kj}) \quad \sum_j{}' P^{\mathrm{h}}_{ij} = P_i - P^{\mathrm{h}}_{il} \quad \sum_i{}'\sum_j{}' P^{\mathrm{h}}_{ij} = 1 - P_k - P_l + P^{\mathrm{h}}_{kl}$$

Met ingevulde deelsommen:

$$\begin{aligned}
P^{\mathrm{h}}(y_{i\lambda}) &= \sum_i{}'(P_i - P^{\mathrm{h}}_{il})y_{i\lambda} = \sum_i{}' P_i y_{i\lambda} - P_l x_\lambda \\
P^{\mathrm{h}}(x_{j\lambda}) &= \sum_j{}'(P_j - P^{\mathrm{h}}_{kj})x_{j\lambda} = \sum_j{}' P_j x_{j\lambda} - P^{\mathrm{h}}_{kl}x_\lambda \\
P^{\mathrm{h}}_\lambda &= \lambda(P^{\mathrm{h}}_{kl}x_\lambda + (\sum_i{}' P_i y_{i\lambda} - P_l x_\lambda) + (\sum_j{}' P_j x_{j\lambda} - P^{\mathrm{h}}_{kl}x_\lambda)) \\
&= \lambda(\sum_i{}' P_i y_{i\lambda} + \sum_j{}' P_j x_{j\lambda} - P_l x_\lambda) \\
&= \lambda(\sum_i{}' P_i y_{i\lambda} + \sum_j{}' P_j y_{j\lambda} - x_\lambda)
\end{aligned}$$

De laatse regel is toegevoegd om de symmetrie tussen i en j te tonen.
De kans op beide eigenvectoren λ en λ' tezamen is

$$\begin{aligned}
P^{\mathrm{h}}_{\lambda,\lambda'} &= \sum_{ij}{}' P^{\mathrm{h}}_{ij}v_{ij\lambda}v_{ij\lambda'} = \sum_{ij}{}' \lambda v_{ij\lambda}v_{ij\lambda'} + \lambda\sum_{ij}{}' P^{\mathrm{h}}_{ij}(y_{i\lambda}+x_{j\lambda})v_{ij\lambda'} \\
&= \lambda\delta_{\lambda,\lambda'} + \lambda\lambda'\sum_{ij}{}'(y_{i\lambda}+x_{j\lambda})(v_{ij\lambda'} + P^{\mathrm{h}}_{ij}(y_{i\lambda'}+x_{j\lambda'})) \\
&= \lambda\delta_{\lambda,\lambda'} + \lambda\lambda'\left(\sum_i{}' y_{i\lambda}v_{i\lambda'} + \sum_j{}' x_{j\lambda}v_{j\lambda'} + P^{\mathrm{h}}((y_{i\lambda}+x_{j\lambda})(y_{i\lambda'}+x_{j\lambda'}))\right)
\end{aligned}$$

met daarin:

$$
\begin{aligned}
\sum_i{}' y_{i\lambda} v_{i\lambda'} &= \sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda} x_{i\lambda'} = \sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda} (y_{i\lambda'} - x_{\lambda'}) = \sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda} y_{i\lambda'} - P_l x_\lambda x_{\lambda'} \\
\sum_j{}' x_{j\lambda} v_{j\lambda'} &= \sum_j{}' P^{\mathrm{h}}_{kj} x_{j\lambda} x_{j\lambda'} \\
P^{\mathrm{h}}(y_{i\lambda} y_{i\lambda'}) &= \sum_i{}' (\sum_j{}' P^{\mathrm{h}}_{ij}) y_{i\lambda} y_{i\lambda'} = \sum_i{}' (P_i - P^{\mathrm{h}}_{il}) y_{i\lambda} y_{i\lambda'} \\
P^{\mathrm{h}}(x_{j\lambda} x_{j\lambda'}) &= \sum_j{}' (P_j - P^{\mathrm{h}}_{kj}) x_{j\lambda} x_{j\lambda'} \\
P^{\mathrm{h}}(y_{i\lambda} x_{j\lambda'}) &= \sum_i{}' \sum_j{}' P^{\mathrm{h}}_{ij} y_{i\lambda} x_{j\lambda'} \qquad P^{\mathrm{h}}(x_{j\lambda} y_{i\lambda'}) = \sum_i{}' \sum_j{}' P^{\mathrm{h}}_{ij} x_{j\lambda} y_{i\lambda'}
\end{aligned}
$$

zodat na substitutie:

$$
\begin{aligned}
P^{\mathrm{h}}_{\lambda,\lambda'} = \lambda\delta_{\lambda,\lambda'} + \lambda\lambda' \Bigg(& \sum_i{}' P^{\mathrm{h}}_{il} y_{i\lambda} y_{i\lambda'} - P_l x_\lambda x_{\lambda'} + \sum_j{}' P^{\mathrm{h}}_{kj} x_{j\lambda} x_{j\lambda'} + \\
& \sum_i{}' (P_i - P^{\mathrm{h}}_{il}) y_{i\lambda} y_{i\lambda'} + \sum_j{}' (P_j - P^{\mathrm{h}}_{kj}) x_{j\lambda} x_{j\lambda'}) + \sum_i{}' \sum_j{}' P^{\mathrm{h}}_{ij} (y_{i\lambda} x_{j\lambda'} + x_{j\lambda} y_{i\lambda'})) \\
= \lambda\delta_{\lambda,\lambda'} + \lambda\lambda' \Bigg(& \sum_i{}' P_i y_{i\lambda} y_{i\lambda'} + \sum_j{}' P_j x_{j\lambda} x_{j\lambda'} - P_l x_\lambda x_{\lambda'} + \sum_{ij}{}' P^{\mathrm{h}}_{ij} (y_{i\lambda} x_{j\lambda'} + x_{j\lambda} y_{i\lambda'}) \Bigg)
\end{aligned}
$$

Als dit vereenvoudigd zou kunnen worden dan zouden we een goed uitganspunt hebben voor de covariantie. Dat zie ik helaas niet in, zodat verder uitwerken van de covariantie niet tot zinvolle resultaten leidt.
De covariantie van beide eigenvectoren λ en λ' is, met substitutie van P^{h}_λ:

$$
\begin{aligned}
\mathbf{covar}^{\mathrm{h}}_{\lambda\lambda'} = & \lambda\delta_{\lambda,\lambda'} + \lambda\lambda' \Bigg(\sum_i{}' P_i y_{i\lambda} y_{i\lambda'} + \sum_j{}' P_j x_{j\lambda} x_{j\lambda'} - P_l x_\lambda x_{\lambda'} + \sum_{ij}{}' P^{\mathrm{h}}_{ij} (x_{j\lambda} y_{i\lambda'} + y_{i\lambda} x_{j\lambda'}) - \\
& - (\sum_i{}' P_i y_{i\lambda} + \sum_j{}' P_j x_{j\lambda} - P_l x_\lambda)(\sum_i{}' P_i y_{i\lambda'} + \sum_j{}' P_j x_{j\lambda'} - P_l x_{\lambda'}) \Bigg) \\
= & \lambda\delta_{\lambda,\lambda'} + \lambda\lambda' \Bigg(\sum_i{}' P_i y_{i\lambda} y_{i\lambda'} - (\sum_i{}' P_i y_{i\lambda})(\sum_i{}' P_i y_{i\lambda'}) + \\
& + \sum_{ij}{}' P^{\mathrm{h}}_{ij} y_{i\lambda} x_{j\lambda'} - (\sum_i{}' P_i y_{i\lambda} - P_l x_\lambda)(\sum_j{}' P_j x_{j\lambda'} - P_l x_{\lambda'}) + \\
& + \sum_{ij}{}' P^{\mathrm{h}}_{ij} x_{j\lambda} y_{i\lambda'} - (\sum_j{}' P_j x_{j\lambda} - P_l x_\lambda)(\sum_i{}' P_i y_{i\lambda'} - P_l x_{\lambda'}) \\
& + \sum_j{}' P_j x_{j\lambda} x_{j\lambda'} - (\sum_j{}' P_j x_{j\lambda})(\sum_j{}' P_j x_{j\lambda'}) - (P_l - P_l^2) x_\lambda x_{\lambda'} \Bigg)
\end{aligned}
$$

De covarianties zijn in het algemeen niet orthonormaal (dan zou alleen de delta-functie overblijven).

De vraag is hoe we de eigenschappen van de nulvectoren $y_{i\lambda}$ van $C_{i\lambda}$, bij eigenwaarde λ, kunnen benutten. We hebben een eigenschap nodig als die tussen P^{h}_{ij} en $A_{ij\lambda}$.

4 Rekenvoorbeelden (met P^{h})

[2017]: **Voorbeeld** Waar het mee is begonnen: een 3x2 tabel.

$$n = \begin{array}{c|cc|c} & G & \bar{G} & \\ Bh & 21 & 19 & 40 \\ Bn & 22 & 19 & 41 \\ \bar{B} & 5 & 11 & 16 \\ \hline & 28 & 49 & 97 \end{array} \qquad P^{\mathrm{h}} = \begin{array}{cc|c} 0{,}2165 & 0{,}1959 & 0{,}4124 \\ 0{,}2268 & 0{,}1959 & 0{,}4227 \\ 0{,}0515 & 0{,}1134 & 0{,}1649 \\ \hline 0{,}4948 & 0{,}5052 & 1{,}0000 \end{array}$$

De covariantietabel en bijpassende correlatietabel worden:

$$\textbf{covar}= \begin{array}{cc|c} 0.01244 & -0.01244 & 0.2423 \\ 0.01765 & -0.01765 & 0.2440 \\ -0.03009 & 0.03009 & 0.1377 \\ \hline 0.2500 & 0.2500 & 1.0000 \end{array} \qquad \textbf{r}= \begin{array}{cc} 0.0505 & -0.0505 \\ 0.0715 & -0.0715 \\ -0.1622 & 0.1622 \end{array}$$

Van de tabel blijft over: i=1, 2, k=3 en j=1, l=2. Twee paren i=1,j=1 en i=2 j=1:

$$P^{\mathrm{h}}_{ij} = \begin{pmatrix} 21/97 \\ 22/97 \end{pmatrix} \quad P^{\mathrm{h}}_{kj} = (5/97) \quad P^{\mathrm{h}}_{il} = \begin{pmatrix} 19/97 \\ 19/97 \end{pmatrix} \quad P^{\mathrm{h}}_{kl} = 11/97$$

$$F = \delta_{ii'}(\delta_{i1}(97/21{+}97/19){+}\delta_{i2}(97/22{+}97/19)){+}(97/5{+}97/11)$$

$$= 97/55 \begin{pmatrix} 8584/19/21 & 16 \\ 16 & 813/38 \end{pmatrix}$$

Hiermee bepalen we $E = F - \delta/\lambda$, en uit determinant nul de eigenwaarden en -vectoren:

$$E = \begin{pmatrix} \frac{832648}{21945} - 1/\lambda & \frac{1552}{55} \\ \frac{1552}{55} & \frac{78861}{2090} - 1/\lambda \end{pmatrix}$$

$$\det E = \frac{1}{\lambda^2} - \frac{3321377}{482790\lambda} + \frac{971422796}{1528835} = \frac{9173010 - 694167793\lambda + 5828536776\lambda^2}{9173010\lambda^2}$$

$$= \frac{5828536776}{9173010\lambda^2}(\lambda^2 - \frac{73777}{619464}\lambda + \frac{1528835}{971422796})$$

$$\lambda = \frac{73777}{1238928} \pm \frac{3971}{120176016}\sqrt{1806361} = \begin{pmatrix} 0.015139 \\ 0.10396 \end{pmatrix} \qquad v_{ij\lambda} = \begin{pmatrix} 0.7084 & 0.7058 \\ -0.7058 & 0.7084 \end{pmatrix}$$

Kansen en covarianties van de nieuwe eigenvectoren blijken aardig orthogonaal:

$$P_\lambda = \begin{pmatrix} 0.3134 & 0.0078 \end{pmatrix} \quad \textbf{covar}_{\lambda\lambda'} = \begin{pmatrix} .1234 & .0027 \\ .0027 & .2216 \end{pmatrix} S_\lambda = \begin{pmatrix} .3513 & .4707 \end{pmatrix}$$

$$\rho_{\lambda\lambda'} = .0162$$

Uit de e_λ en λ volgt χ^2:

$$e = \begin{pmatrix} e_{\lambda 1} & e_{\lambda 2} \end{pmatrix} = \begin{pmatrix} S_{\lambda 1}(1{-}\rho_{\lambda 1}) & S_{\lambda 2}(1{-}\rho_{\lambda 2}) \end{pmatrix}$$

$$= \begin{pmatrix} 0,3513(1{-}\rho_{\lambda 1}) & 0,4707(1{-}\rho_{\lambda 2}) \end{pmatrix}$$

$$\chi^2/N = \Sigma_\lambda \frac{e^2_\lambda}{\lambda} = \frac{e^2_{\lambda 1}}{\lambda_1} + \frac{e^2_{\lambda 2}}{\lambda_2} = \frac{e^2_{\lambda 1}}{0,0151} + \frac{e^2_{\lambda 2}}{0,1040} =$$

$$= \frac{0.1234(1{-}\rho_{\lambda 1})^2}{0.01513} + \frac{0.2216(1{-}\rho_{\lambda 2})^2}{0.1040} = 8.17(1{-}\rho_{\lambda 1})^2 + 2.13(1{-}\rho_{\lambda 2})^2$$

Aan de andere kant volgen de e_{ij} uit de e_λ, en dus ook χ^2:

$$e_{12} = -e_{11} \quad e_{22} = -e_{21} \quad e_{31} = -(e_{11} + e_{21}) = 0,0026e_{\lambda 1} - 1,4112e_{\lambda 2} = -e_{32}$$

$$\chi^2/N = \Sigma_{ij}\frac{e^2_{ij}}{P^{\mathrm{h}}_{ij}} = (\frac{1}{P^{\mathrm{h}}_{11}}{+}\frac{1}{P^{\mathrm{h}}_{12}}{+}\frac{1}{P^{\mathrm{h}}_{31}}{+}\frac{1}{P^{\mathrm{h}}_{32}})e^2_{11} + (\frac{1}{P^{\mathrm{h}}_{21}}{+}\frac{1}{P^{\mathrm{h}}_{22}}{+}\frac{1}{P^{\mathrm{h}}_{31}}{+}\frac{1}{P^{\mathrm{h}}_{32}})e^2_{21}+$$

$$+2(\frac{1}{P^{\mathrm{h}}_{31}}{+}\frac{1}{P^{\mathrm{h}}_{32}})e_{11}e_{21} = 37,9593e^2_{11} + 56,4716e_{11}e_{21} + 37,7496e^2_{21}$$

De kwadratische vorm wordt inderdaad gediagonalizeerd door e_{ij} uit te drukken in e_λ:

$$e_{ij} = \Sigma_\lambda v_{ij\lambda} e_\lambda = \begin{pmatrix} .7084e_{\lambda 1} - .7058e_{\lambda 2} & .7058e_{\lambda 1} + .7084e_{\lambda 2} \end{pmatrix}$$

$$\chi^2/N = 37,9593e_{11}^2 + 56,4716e_{11}e_{21} + 37,7496e_{21}^2 = 66,0904e_{\lambda_1}^2 + 9,6185e_{\lambda_2}^2$$

Rechtstreeks toegepast op e_{11}^2 alleen, blijkt die iets meer bepaald door 'λ_2':

$$e_{11}^2 = (.7084e_{\lambda_1} - .7058e_{\lambda_2})^2 = 0,5018e_{\lambda_1}^2 - 1,0e_{\lambda_1}e_{\lambda_2} + 0,4982e_{\lambda_2}^2$$
$$= 0,06192(1-\rho_{\lambda_1})^2 - 0,1577(1-\rho_{\lambda_1})(1-\rho_{\lambda_2}) + 0,1103(1-\rho_{\lambda_2})^2$$
$$= 0,1686\left(0,8041(1-\rho_{\lambda_2}) - 0,5944(1-\rho_{\lambda_1})\right)^2 + 0,0036(0,8041(1-\rho_{\lambda_1}) + 0,5944(1-\rho_{\lambda_2}))^2$$
$$= 0,2500.0,2423(0,0505 - \rho_{11})^2 = 0,06058(0,0505 - \rho_{11})^2$$

De matrices A en vectoren B zijn functies van λ:

$$A_{\lambda ij} = \begin{pmatrix} 1/(1/\lambda - 97/21) \\ 1/(1/\lambda - 97/22) \end{pmatrix} = \lambda \begin{pmatrix} 21/(21-97\lambda) \\ 22/(22-97\lambda) \end{pmatrix}$$
$$A_{i\lambda} = A_{ij\lambda} \quad A_{j\lambda} = \lambda \begin{pmatrix} 21/(21-97\lambda) \\ 22/(22-97\lambda) \end{pmatrix} =$$
$$A_{\lambda j} = \lambda\left((924 - 4171\lambda)/(21-97\lambda)(22-97\lambda)\right)$$
$$A_\lambda = A_{j\lambda} = \lambda\left((924 - 4171\lambda)/(21-97\lambda)(22-97\lambda)\right)$$
$$B_{\lambda i} = \begin{pmatrix} 19/97 \\ 19/97 \end{pmatrix} - \lambda \begin{pmatrix} 21/(21-97\lambda) \\ 22/(22-97\lambda) \end{pmatrix} = 1/97 \begin{pmatrix} (399 - 3880\lambda)/(21-97\lambda) \\ (418 - 3977\lambda)/(22-97\lambda) \end{pmatrix}$$
$$B_{j\lambda} = 5/97 - \lambda\left((924 - 4171\lambda)/(21-97\lambda)(22-97\lambda)\right) =$$
$$B_{j\lambda} = 1/97\left((2310 + 451613\lambda - 110483\lambda^2)/(21-97\lambda)(22-97\lambda)\right)$$

De matrices $\overline{B_\lambda}$ en C kunnen we nu ook bepalen:

$$\overline{B_{i\lambda}} = 1/97 \begin{pmatrix} \frac{399-3880\lambda}{21-97\lambda} & 0 \\ 0 & \frac{418-3977\lambda}{22-97\lambda} \end{pmatrix} -$$
$$-\lambda \begin{pmatrix} \frac{21}{21-97\lambda} \\ \frac{22}{22-97\lambda} \end{pmatrix} \left(\frac{(21-97\lambda)(22-97\lambda)}{2310+451613\lambda-110483\lambda^2}\right) \begin{pmatrix} \frac{21}{21-97\lambda} & \frac{22}{22-97\lambda} \end{pmatrix} \lambda =$$
$$\overline{B_{i\lambda}} = \frac{1/97}{2310+451632\lambda-110483\lambda^2} \begin{pmatrix} 43890-2323247\lambda+13915911\lambda^2 & -97*44814\lambda^2 \\ -97*44814\lambda^2 & 43890-2323247\lambda+13962956\lambda^2 \end{pmatrix}$$
$$C_i = \frac{1/47531}{2310+451632\lambda-110483\lambda^2} \begin{pmatrix} 3(172946829\lambda^2-24651580\lambda+438900) & -(376040094\lambda^2-39884363\lambda+833910) \\ -(376040094\lambda^2-39884363\lambda+833910) & 4(130286423\lambda^2-18488685\lambda+329175) \end{pmatrix}$$

Met behulp van C_λ zijn alternatief de eigenwaarden en vectoren te bepalen. Daartoe wordt eerst de determinant van $C_{i\lambda}$ bepaald, en vervolgens de eigenwaarden als de oplossingen van $|C_{i\lambda}| = |E|/|D_1||D_l| = 0$:

$$\mathbf{det}(C_{i\lambda}) = \frac{5828536776\lambda^2 - 694167793\lambda + 9173010}{461041(2310 - 110483\lambda^2 + 451632\lambda)}$$
$$0 = 5828536776\lambda^2 - 694167793\lambda + 9173010$$
$$\lambda = \frac{73777}{1238928} \pm \frac{3971}{120176016}\sqrt{1806361} = \begin{pmatrix} 0.015139 \\ 0.10396 \end{pmatrix}$$

De matrices $\overline{B_{i\lambda}}^{-1}$ tenslotte door omkeren:

$$\overline{B_{i\lambda}}^{-1} = \frac{97^2}{388394111(\lambda - \frac{498959}{4004063})(\lambda + \frac{119130}{55484873})} \begin{pmatrix} \frac{43890-2323247\lambda+13962956\lambda^2}{97} & 44814\lambda^2 \\ 44814\lambda^2 & \frac{43890-2323247\lambda+13915911\lambda^2}{97} \end{pmatrix}$$
$$1 - P_{il}^{\mathrm{h}} \overline{B_{\lambda i}}^{-1} P_{il}^{\mathrm{h}}/P_l = \frac{1}{49} \frac{9173010 - 694167793\lambda + 5828536776\lambda^2}{833910 - 48399023\lambda + 388394111\lambda^2}$$

Merk op, dat $1-P^{\mathrm{h}}_{il}\overline{B_{\lambda i}}^{-1}P^{\mathrm{h}}_{il}/P_l$ evenredig is met $|E|$, dus dezelfde eigenwaarden geeft, en het omgekeerde van $1+{P^{\mathrm{h}}_{il}}^t C^{-1}_{i\lambda}P^{\mathrm{h}}_{il}/P_l$:

$$C^{-1}_{i\lambda} = \frac{1}{5828536776\lambda^2-694167793\lambda+9173010}\begin{pmatrix} 388(130286423\lambda^2-18488685\lambda+329175) & 97(376040094\lambda^2+833910-39884363\lambda) \\ 97(376040094\lambda^2+833910-39884363\lambda) & 291(172946829\lambda^2-24651580\lambda+438900) \end{pmatrix}$$

$$1+{P^{\mathrm{h}}_{il}}^t C^{-1}_{i\lambda}P^{\mathrm{h}}_{il}/P_l = 49\frac{388394111\lambda^2-48399023\lambda+833910}{5828536776\lambda^2-694167793\lambda+9173010}$$

Voorbeeld Neem de volgende 3x3 tabel met covariantie en correlatieberekening

$$P^{\mathrm{h}}_{ij} = \begin{pmatrix} 0.02 & 0.03 & 0.05 \\ 0.07 & 0.11 & 0.12 \\ 0.17 & 0.20 & 0.23 \end{pmatrix} \quad P^{\mathrm{h}}_{kj} = \begin{pmatrix} 0.17 & 0.20 \end{pmatrix} \quad P^{\mathrm{h}}_{il} = \begin{pmatrix} 0.05 \\ 0.12 \end{pmatrix} \quad P^{\mathrm{h}}_{kl} = 0.23$$

$$P_{ij} = \left(\begin{array}{cc|cc} 0.02 & 0.03 & 0.05 & 0.10 \\ 0.07 & 0.11 & 0.12 & 0.30 \\ \hline 0.17 & 0.20 & 0.23 & 0.60 \\ 0.26 & 0.34 & 0.40 & 1.00 \end{array}\right) \ \mathbf{covar} = \left(\begin{array}{ccc|c} -0.006 & -0.004 & 0.01 & .09 \\ -0.008 & 0.008 & 0 & .21 \\ 0.014 & -0.004 & -0.01 & .24 \\ \hline .1924 & .2244 & .24 & 1. \end{array}\right)$$

$$\rho = \begin{pmatrix} -0.0456 & -0.0281 & 0.0680 \\ -0.0398 & 0.0369 & 0.0000 \\ 0.0652 & -0.0173 & -0.0417 \end{pmatrix}$$

De 2x2 χ^2-matrix F wordt opgebouwd uit 2x2 ondermatrices D_j:

$$D_1 = \begin{pmatrix} 1/0.02+1/0.17 & 1/0.17 \\ 1/0.17 & 1/0.07+1/0.17 \end{pmatrix} = \begin{pmatrix} 55.88 & 5.882 \\ 5.882 & 20.17 \end{pmatrix}$$

$$D_2 = \begin{pmatrix} 1/0.03+1/0.20 & 1/0.20 \\ 1/0.20 & 1/0.11+1/0.20 \end{pmatrix} = \begin{pmatrix} 38.33 & 5.00 \\ 5.00 & 14.09 \end{pmatrix}$$

$$D_3 = D_l = \begin{pmatrix} 1/0.05+1/0.23 & 1/0.23 \\ 1/0.23 & 1/0.12+1/0.23 \end{pmatrix} = \begin{pmatrix} 24.35 & 4.348 \\ 4.348 & 12.68 \end{pmatrix}$$

$$F = \begin{pmatrix} D_1+D_3 & D_3 \\ D_3 & D_2+D_3 \end{pmatrix} = \begin{pmatrix} 80.23 & 10.23 & 24.35 & 4.35 \\ 10.23 & 32.85 & 4.35 & 12.68 \\ 24.35 & 4.35 & 62.68 & 9.35 \\ 4.35 & 12.68 & 9.35 & 26.77 \end{pmatrix}$$

De eigenwaarden en bijbehorende orthonormale eigenvectoren zijn (per rij):

$$\lambda = \begin{pmatrix} 0.009925 \\ 0.02136 \\ 0.02550 \\ 0.06346 \end{pmatrix} \quad v_{ij\lambda} = \begin{pmatrix} .79150 & .18318 & .56367 & .14913 \\ -.59573 & .021254 & .76574 & .24145 \\ -.093090 & .77880 & -.27012 & .55843 \\ .099786 & -.59955 & -.15151 & .77950 \end{pmatrix}$$

De eigenwaarden van matrix F zijn: 100.76; 46.81; 39.21; 15.75, dat zijn onze omgekeerde eigenwaarden $1/\lambda$. De kleinste $\lambda = 0.0099 < 0.02$ is kleiner dan de kleinste kans P_{ij} uit de 2x2 submatrix, de grootste $\lambda = 0.06346 < 0.11$ is kleiner dan de grootste kans uit de submatrix. Merk op dat de eigenvectoren sterk geconcentreerd zijn op een enkel variabelen paar: de grootste eigenwaarde op $i = 1, j = 1$, de op een na grootste op $i = 1, j = 2$, de op een na kleinste op $i = 2, j = 1$ en de kleinste op $i = 2, j = 2$.
Tenslotte χ^2 uitgedrukt in eigenwaarden λ en excessvariabelen e_λ met $\mathbf{cov}^h_\lambda = \Sigma ij'\mathbf{cov}^h_{ij}v_{ij\lambda}$:

$$\mathbf{cov}^h_\lambda = \begin{pmatrix} -.00727 & .00227 & -.00012 & .01104 \end{pmatrix}$$
$$\chi^2/N = 100.8(c_1+.0073)^2 + 46.81(c_2-.0023)^2 + 2.662(c_3+.0001)^2 + 1.437(c_4-.011)^2$$

De eigenvectoren blijken statistisch *enigzins* afhankelijk te zijn; met de kansen $P_\lambda = \Sigma'_{ij}P^{\mathrm{h}}_{ij}v_{\lambda ij}$, covarianties $\Sigma'_{ij} P^{\mathrm{h}}_{ij} v_{\lambda ij} v_{\lambda' ij} - P_\lambda P_{\lambda'}$ en correlaties $\rho = \mathbf{cov}/SS$:

$$P_\lambda = \begin{pmatrix} .061967 & .039104 & .10598 & .041227 \end{pmatrix}$$

$$\mathbf{cov}_{\lambda\lambda'} = \begin{pmatrix} .02302 & .00533 & .00654 & .00156 \\ & .02960 & .00675 & .01353 \\ & & .06789 & .01187 \\ & & & .09119 \\ \hline .15171 & .17206 & .26055 & .30198 \end{pmatrix} \quad \rho_{\lambda\lambda'} = \begin{pmatrix} 1 & .20413 & .16541 & .03410 \\ & 1 & .15056 & .26040 \\ & & 1 & .15086 \\ & & & 1 \end{pmatrix}$$

In dit voorbeeld zijn de standaardafwijkingen veel groter dan de λ-waarden. Onafhankelijkheid van de gemengde variabelen kan worden verkregen door de covariantie te diagonaliseren. Door de draaing van het assenstelsel ontstaan nieuwe gemengde variabelen aangeduid met μ, bestaande uit combinaties van eigenvectoren v_λ, die wel onafhankelijk zijn met standaardafwijking

$$S_\mu = \begin{pmatrix} .1394 & .1698 & .2523 & .3159 \end{pmatrix}$$

Merk op, dat de standaardafwijkingen door de diagonalisatie zowel groter als kleiner zijn geworden als verwacht kon worden. In deze situatie is χ^2 echter niet meer minimaal.

De submatrices $D_{j\lambda}$ zijn:

$$D_{1\lambda} = \begin{pmatrix} 950\lambda - 17 & 100x\lambda \\ 100\lambda & 2400\lambda/7 - 17 \end{pmatrix} /17\lambda \quad \det(D_{1\lambda}) = \frac{119 - 9050\lambda + 130000\lambda^2}{119\lambda^2} = \frac{p_1}{119\lambda^2}$$

$$D_{2\lambda} = \begin{pmatrix} 115\lambda/3 - 1 & 5\lambda \\ 5\lambda & 155\lambda/11 - 1 \end{pmatrix} /\lambda \quad \det(D_{2\lambda}) = \frac{33 - 1730\lambda + 17000\lambda^2}{33\lambda^2} = \frac{p_2}{33\lambda^2}$$

De matrix $C_{i\lambda} = \Sigma_j D^{-1}_{j\lambda}$:

$$C_{i\lambda} = \begin{pmatrix} 250p_3 & -20p_4 \\ -20p_4 & 16p_5 \end{pmatrix} /4000p_1p_2$$

$$p_3 = 31670000000\lambda^4 - 4998970000\lambda^3 + 256035300\lambda^2 - 4788280\lambda + 27489$$

$$p_4 = 11781 - 1513560\lambda + 74455500\lambda^2 - 1677100000\lambda^3 + 13300000000\lambda^4$$

$$p_5 = 115785000000\lambda^4 - 15270187500\lambda^3 + 679985750x\lambda^2 - 12558420\lambda + 82467$$

$$\det(C_{i\lambda}) = \frac{789597500000\lambda^4 - 94950591250\lambda^3 + 3643218775\lambda^2 - 54878850\lambda + 270963}{20000p_1p_2}$$

5 Naschrift

5.1 De significantie volgens Pearson en Fisher

2008: We gaan op zoek naar de theorie van de toevalsaantallen tabel. De klassieke bron (volgens Wikipedia) is K.Pearson, On the criterion that a given sytem of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Phil. Mag., **1**_(1900), p. 157._ Deze behandelt de aantallen als onafhankelijke variabelen die de normale verdeling hebben, met als enige beperking dat de som van alle afwijkingen nul is. De normale verdeling wordt bepaald door chi-kwadraat die afhangt van gemiddelde μ en spreiding σ, of te wel de variantie $\mathbf{var} = \sigma^2$:

$$\begin{aligned} f &\propto e^{-\frac{1}{2}\chi^2} \\ \chi^2 &= (n-\mu)^2/\mathbf{var}(n), \end{aligned}$$

Bij de toevalstabel neemt Pearson voor de variantie het gemiddelde $\mathbf{var} = \mu$. In een tabel met onderling onafhankelijke cellen maar met constant totaal aantal N blijft de variantie echter onder het gemiddelde:

$$\begin{aligned} \mathbf{E}(n) &= \mu = PN \\ \mathbf{var}(n) &= PN(1-P) = \mu(1-P) = \mu - \mu P < \mu \end{aligned}$$

R.A.Fisher, ON THE INTERPRETATION OF χ^2 FROM CONTINGENCY TABLES, AND THE CALCULATION OF P, Journal of the Royal Statistical Society, *Vol.***85***, No. 1.(Jan., 1922), pp. 87–94*, heeft laten zien dat het laten fluctueren van het totale aantal leidt tot de gevraagde gelijkheid. Omdat het totale aantal een Poissonverdeling heeft is $\mathbf{var}(N) = N$; de variantie van het aantal wordt door de fluctuatie van N vergroot met de term: $\mathbf{var}(PN) = P^2\mathbf{var}(N) = P^2N = \mu P$, juist het ontbrekende deel.

Hij (Fisher) wijst er in een voetnoot op dat de aantallen in de tabel op te vatten zijn als onafhankelijke variabelen die alle de Poisson verdeling hebben, maar met beperkingen voor de somaantallen van rijen en kolommen. De Poissonverdeling wordt bepaald door het gemiddelde alleen:

$$\begin{aligned} f &\propto \mu^n/n! \\ \mathbf{E}(n) &= \mu \qquad \mathbf{var}(n) = \mu, \end{aligned}$$

waardoor $\mathbf{var} = \mu$ exact is.

Bij grotere totale aantallen, zoals $N \geq 10$, maar bij k cellen vooral $N^2/k \geq 10$, is de normale verdeling als benadering bruikbaar, en dus het gebruik van de chi-kwadraat methode. Zie bijvoorbeeld het werk van Kenneth J. Koehler, Kinley Larntz, AN EMPIRICAL INVESTIGATION OF GOODNESS-OF-FIT STATISTICS FOR SPARSE MULTINOMIALS, Journal of the American Statistical Association, *Vol.***75***, No. 370. (June, 1980), pp. 336-344.* Zij onderzoeken met de MonteCarlo methode (computer testen) de betrouwbaarheid van het gebruik van 1) de chi-kwadraat methode en 2) de likelihood ratio methode. Volgens hen is de laatste iets beter bij kleine frekwenties, en sneller te benaderen met de normale verdeling, maar zij geven aan dat het gebruik van chi-kwadraat voldoende goed is: "the chi-squared approximation for the Pearson statistic is quite adequate at the .05 and .01 nominal levels for expected frequencies as low as .25 when $k \geq 3$, $n \geq 10$, $n^2/k \geq 10$". Daarbij is de variabele k het aantal cellen (bij ons het product van aantal rijen en aantal kolommen), de variabele n het totale aantal (bij ons N, de 'sample size'), en 'frequency' het aantal in de cel (bij ons n). Maar er moeten niet te veel verwachte aantallen onder de 1 zijn: "The chi-squared approximation for the Pearson statistic produces inflated rejection levels for unsymmetrical null hypotheses that contain many expected frequencies smaller than one."

5.2 Canonieke variabelen volgens Hotelling.

2015: Naar aanleiding van vragen over het gebruik van correlaties, kwam Egbert een artikel op het spoor over canonieke transformaties van variabelen via de matrix van correlaties: Harald Hotelling, RELATIONS BETWEEN TWO SETS OF VARIATES, Biometrika, Vol. 28, No. 3/4 (Dec., 1936), pp. 321-377. De ene set bestaat uit 'oorzaak'variabelen, de andere uit 'gevolg'variabelen. De naam 'oorzaak' of 'gevolg' heeft geen causale betekenis: ze zijn verwisselbaar. Eigenvectoren van de correlatiematrix van oorzaakvariabelen worden vergeleken met eigenvectoren van de correlatiematrix van gevolgvariabelen, volgens het criterium dat de oorzaak-gevolgcorrelatie ρ maximaal is. Een niet-negatieve ρ wordt een *canonieke correlatie* genoemd, de bijpassende samenstelling van eigenvectoren *canonieke variabelen.* Canonieke variabelen zijn onderling onafhankelijk, uitgezonderd de bijpassende. De correlatiematrix van de canonieke variabelen is bijna diagonaal: het oorzaak en gevolg deel is diagonaal 1, de onderlinge oorzaak-gevolg delen zijn diagonaal, voorzover mogelijk, met daarop de ρ's.

Hotelling beschouwt twee maten voor de afhankelijkheid of onafhankelijk.

1. Het product Q van de verschillende canonieke waarden ρ;
2. Het product Z van de verschillende canonieke waarden van $(1-\rho^2)$ (op het teken na de determinant van de canonieke correlatiematrix);

In het algemeen geldt $0 \leq Z \leq (1-Q)^2 \leq (1-Q^2) \leq 1$ met gelijkheid alleen als $Z = 0$, $Q = 1$ (oorzaak gevolg gelijk) of $Z = 1$, $Q = 0$ (oorzaak onafhankelijk gevolg). Verder onderzocht hij het asymptotisch gedrag van steekproeven en concludeerde dat de steekproefspreiding gaat als $1/\sqrt{n}$:

- van ρ_i als $\sqrt{1-\rho_1^2}/\sqrt{n}$;
- van Q als $Q\sqrt{\sum_i(1-\rho_i^2)}/\sqrt{n}$;

- van Z als $2Z\sqrt{\sum_i \rho_i^2}\sqrt{n}$;

Voor de asymptotische frequentieverdeling heeft hij deelresultaten, en wel voor het geval "twee 'oorzaak' variabelen onafhankelijkheid van 'gevolg' variabelen". Onder de aanname van normale verdeling kan het steekproefgedrag getoetst worden.

Het is duidelijk dat Hotelling een belangrijk probleem heeft opgelost. Neem bijvoorbeeld het 'onafhankelijke geval', waarbij χ^2 de som over alle correlaties is—daar lijkt Hotelling een stap verder te gaan door deze te diagonalizeren naar de canonieke correlaties. Verder behandelt hij in een moeite door ook het 'afhankelijke geval'—en daarvoor had ik nog niet een doorzichtige formule kunnen afleiden! Overigens spelen in mijn aanpak de interne correlaties geen enkele rol (of het moet een verborgen rol zijn), maar alleen oorzaak-gevolg kansen en covarianties, waardoor de resultaten lijken te verschillen.

A Uitwerking: de som der omgekeerde kansen

2005: Voor de som der omgekeerde th(eoretische) geldt:

$$\text{som omgekeerden} = \frac{1}{P_{BG}^{\text{th}}} + \frac{1}{P_{\bar{B}G}^{\text{th}}} + \cdots = \frac{P_{BG}^{\text{th}} + P_{\bar{B}G}^{\text{th}}}{P_{BG}^{\text{th}} P_{\bar{B}G}^{\text{th}}} + \cdots$$

De theoretische kans wordt uitgedrukt in de correlatie:

$$P_{BG}^{\text{th}} = P_B P_G + \rho_{B,G} S_B S_G$$

en vergelijkbare uitdrukkingen voor de andere gevallen, waarbij door negatie de correlatie van teken wisselt: $\rho = \rho_{BG} = -\rho_{\bar{B}G}$. Verder geldt:

$$P_{BG}^{\text{th}} + P_{\bar{B}G}^{\text{th}} = P_B P_G + \rho_{B,G} S_B S_G + P_{\bar{B}} P_G - \rho_{B,G} S_B S_G = P_G$$

De *som der omgekeerden* wordt daarmee:

$$\text{som omgekeerden} \quad = \quad \frac{P_G}{P_{BG}^{\text{th}} P_{\bar{B}G}^{\text{th}}} + \frac{P_{\bar{G}}}{P_{\bar{B}\bar{G}}^{\text{th}} P_{B\bar{G}}^{\text{th}}} = \frac{P_G P_{\bar{B}\bar{G}}^{\text{th}} P_{B\bar{G}}^{\text{th}} + P_{\bar{G}} P_{BG}^{\text{th}} P_{\bar{B}G}^{\text{th}}}{P_{BG}^{\text{th}} P_{\bar{B}\bar{G}}^{\text{th}} P_{\bar{B}G}^{\text{th}} P_{B\bar{G}}^{\text{th}}}$$

Uitwerking van de *teller* naar $\rho S_B S_G$ geeft:

$$\begin{aligned} &= \quad P_G(P_{\bar{B}} P_{\bar{G}} + \rho S_B S_G)(P_B P_{\bar{G}} - \rho S_B S_G) + P_{\bar{G}}(P_B P_G + \rho S_B S_G)(P_{\bar{B}} P_G - \rho S_B S_G) \\ &= \quad (P_G + P_{\bar{G}}) P_B P_{\bar{B}} P_G P_{\bar{G}} - (P_G + P_{\bar{G}})(\rho S_B S_G)^2 = (1-\rho^2) S_B^2 S_G^2 \end{aligned}$$

De *noemer* heeft 4 factoren, die $S_B S_G$ gemeenschappelijk hebben:

$$(\rho + \frac{P_B}{S_B}\frac{P_G}{S_G})(\rho + \frac{P_{\bar{B}}}{S_B}\frac{P_{\bar{G}}}{S_G})(\rho - \frac{P_{\bar{B}}}{S_B}\frac{P_G}{S_G})(\rho - \frac{P_B}{S_B}\frac{P_{\bar{G}}}{S_G})(S_B S_G)^4$$

De vierdegraads veelterm in ρ heeft 4 nulpunten, de eerste twee negatief, de volgende twee positief. Het paarsgewijze product van de wortels is 1, zodat de ene (absoluut) kleiner dan 1 en de andere groter dan 1 is. Voor ons is alleen interessant het positieve deel van de veelterm rond 0, tussen het kleinste negatieve en het kleinste positieve nulpunt met een maximum in de buurt van 0. Elk der factoren in de noemer, en chi-kwadraat, is positief tussen de kleinste nulpunten van de noemer: de *minimum correlatie* en de *maximum correlatie*

$$\rho_{\min} = -\min(\frac{P_B}{S_B}\frac{P_G}{S_G}, 1/\frac{P_B}{S_B}\frac{P_G}{S_G}) \quad \rho_{\max} = \min(\frac{P_G}{S_G}/\frac{P_B}{S_B}, 1/(\frac{P_G}{S_G}/\frac{P_B}{S_B}))$$

In het voorbeeld $P_B/S_B = 0,8351/0,3711 = 2,2504 = 1/0.4444$, $P_G/S_G = 0,4948/0,5000 = 0,9897$, zodat $-0,45 \le \rho, r = 0,16 \le 0,44$.

De noemer/$(S_B S_G)^4$ is een veelterm in ρ bestaande uit tweemaal het product van twee factoren van het type $\rho - x$ en $\rho - 1/x$ waarin x een nulpunt is. Het product $(\rho - x)(\rho - 1/x) = \rho^2 - (x + 1/x)\rho + 1$, waarbij de som $|x + 1/x| \ge 2$. Bij de eerste groep van twee factoren is die som $-a = \frac{P_B}{S_B}\frac{P_G}{S_G} + \frac{P_{\bar{B}}}{S_B}\frac{P_{\bar{G}}}{S_G} = (P_B P_G + P_{\bar{B}} P_{\bar{G}})/S_B S_G$. Bij de tweede tweede groep is die som $b = \frac{P_B}{S_B}\frac{P_{\bar{G}}}{S_G} + \frac{P_{\bar{B}}}{S_B}\frac{P_G}{S_G} = (P_B P_{\bar{G}} + P_{\bar{B}} P_G)/S_B S_G$.
De noemer/$(S_B S_G)^4$ wordt daarmee:

$$\begin{aligned} &= (\rho^2 - a\rho + 1)(\rho^2 - b\rho + 1) = (1+\rho^2 - a\rho)(1+\rho^2 - b\rho) \\ &= ((1+\rho^2)^2 - (b+a)\rho(1+\rho^2) + ab\rho^2) \\ &= ((1-\rho^2)^2 + 4\rho^2 - (b+a)\rho((1-\rho)^2 + 2\rho) + ab\rho^2) \\ &= ((1-\rho^2)^2 - (b+a)\rho(1-\rho)^2 + (4 - 2(b+a) + ab)\rho^2) \end{aligned}$$

Uitwerking van $(b+a)$ en $-ab$ naar verschil kansen :

$$
\begin{aligned}
(b+a)S_BS_G &= P_BP_{\bar G} + P_{\bar B}P_G - (P_BP_G + P_{\bar B}P_{\bar G}) = -(P_B - P_{\bar B})(P_G - P_{\bar G}) \\
-abS_B^2S_G^2 &= (P_BP_G + P_{\bar B}P_{\bar G})(P_BP_{\bar G} + P_{\bar B}P_G) = (P_B^2 + P_{\bar B}^2)S_G^2 + (P_G^2 + P_{\bar G}^2)S_B^2 \\
&= ((P_B - P_{\bar B})^2 + 2S_B^2)S_G^2 + ((P_G - P_{\bar G})^2 + 2S_G^2)S_B^2 \\
&= \left((\frac{P_B - P_{\bar B}}{S_B})^2 + (\frac{P_G - P_{\bar G}}{S_G})^2 + 4 \right) S_B^2S_G^2 \Rightarrow \\
4 - 2(b+a) + ab &= 4 + 2\frac{P_B - P_{\bar B}}{S_B}\frac{P_G - P_{\bar G}}{S_G} - ((\frac{P_B - P_{\bar B}}{S_B})^2 + (\frac{P_G - P_{\bar G}}{S_G})^2 + 4) \\
&= -\left(\frac{P_B - P_{\bar B}}{S_B} - \frac{P_G - P_{\bar G}}{S_G} \right)^2
\end{aligned}
$$

geeft voor de noemer$/(S_BS_G)^4$:

$$
(1-\rho^2)^2 + \rho(1-\rho)^2 \frac{P_B - P_{\bar B}}{S_B}\frac{P_G - P_{\bar G}}{S_G} - \rho^2 \left(\frac{P_B - P_{\bar B}}{S_B} - \frac{P_G - P_{\bar G}}{S_G} \right)^2
$$

Voor kansen geldt algemeen dat $P - \bar P = P - (1-P) = 2(P - 1/2)$ tussen -1 en 1 is. Voor $S^2 = P\bar P = P(1-P) = 1/4 - (P-1/2)^2 = 1/4(1 - (P - \bar P)^2)$. We zien dat $|P - \bar P| \le 1$ en $S \le 1/2$. Dus de factor $S_BS_G \le 1/4$, waardoor de rechter kwadratische term $\ge 4\rho^2$. Verder is $(P - \bar P)^2/S^2 = 4(P - \bar P)^2/(1 - (P - \bar P)^2)$ onbeperkt aan de randen $P = 0$ en $P = 1$, maar 0 in het midden $P = 1/2$.

B Begrenzing van de correlatie

2008: De begrenzing van de produktkans begrenst de correlatie:

$$
\begin{aligned}
0 \le P_{BG}^{\text{th}} = P_{BG} = P_BP_G + \rho_{B,G}S_BS_G &\le \min(P_B, P_G) \\
-P_BP_G \le \rho_{B,G}S_BS_G &\le P_BP_{\bar G}, P_GP_{\bar B} \\
-\frac{P_B}{S_B}\frac{P_G}{S_G} \le \rho_{B,G} \le \frac{P_B}{S_B}\frac{P_{\bar G}}{S_G}, \frac{P_G}{S_G}\frac{P_{\bar B}}{S_B} &= \frac{P_B}{S_B}/\frac{P_G}{S_G}, \frac{P_G}{S_G}/\frac{P_B}{S_B} \\
-\frac{P_B}{S_B}\frac{P_G}{S_G} \le \rho_{B,G} &\le \min(\frac{P_B}{S_B}, \frac{P_G}{S_G})/\max(\frac{P_B}{S_B}, \frac{P_G}{S_G})
\end{aligned}
$$

De ondergrens wordt scherper met de produktkans van de complementen:

$$
\begin{aligned}
0 \le P_{\bar B\bar G}^{\text{th}} = P_{\bar B\bar G} = P_{\bar B}P_{\bar G} + \rho_{B,G}S_BS_G &\le P_{\bar B}, P_{\bar G} \\
-1/\frac{P_B}{S_B}\frac{P_G}{S_G} \le \rho_{B,G} &\le \min(\frac{P_B}{S_B}, \frac{P_G}{S_G})/\max(\frac{P_B}{S_B}, \frac{P_G}{S_G})
\end{aligned}
$$

Samengevat is de begrenzing van de correlatie:

$$
\boxed{-\min(\frac{P_B}{S_B}\frac{P_G}{S_G}, 1/\frac{P_B}{S_B}\frac{P_G}{S_G}) \le \rho_{B,G} \le \min(\frac{P_B}{S_B}, \frac{P_G}{S_G})/\max(\frac{P_B}{S_B}, \frac{P_G}{S_G})}
$$

De bovengrens 1 van ρ wordt slechts dan bereikt als de fracties gelijk zijn: $P_B/S_B = P_G/S_G$; maar fractie-kwadraat $P^2/S^2 = P/\bar P = P/(1-P)$, dus $P_B/(1-P_B) = P_G/(1-P_G)$, of $P_B = P_G$. De bovengrens wordt slechts dan bereikt als de kansen gelijk zijn: $P_B = P_G$.

De ondergrens -1 van ρ wordt slechts dan bereikt als het product van de fracties 1 is, of $P_BP_G = S_BS_G$, of $P_BP_G = P_{\bar B}P_{\bar G} = 1 - P_B - P_G + P_BP_G$, of $P_B + P_G = 1$. De ondergrens wordt slechts dan bereikt als de kansen complementair zijn: $P_B + P_G = 1$.

C Eigenschappen van Schur-complementen.

2019: Op grond van het artikel: "On deriving the inverse of a sum of matrices", H. V. Henderson en S. R. Searle, Biometric unit Series BU-647-M, January 1980 (gevonden op internet), geldt voor vierkante niet singuliere matrices A en $A+BDC$, dat de omgekeerden bestaan en dat achtereenvolgens gelijk zijn:

$$1-(A{+}BDC)^{-1}A=(A{+}BDC)^{-1}BDC=A^{-1}BDC(A{+}BDC)^{-1}=$$
$$=A^{-1}BD(1{+}CA^{-1}BD)^{-1}C=A^{-1}B(1{+}DCA^{-1}B)^{-1}DC$$

De laatste twee zijn gelijk als ook D niet singulier is, zodat dan de transformatie $D\to -D^{-1}$ mogelijk is. Als A, D èn $A{-}BD^{-1}C$ niet singulier zijn, dan is ook $D{-}CA^{-1}B$ niet singulier en geldt:

$$A(A{-}BD^{-1}C)^{-1}-1=A(A{-}BD^{-1}C)^{-1}BD^{-1}CA^{-1}=$$
$$=B(D{-}CA^{-1}B)^{-1}CA^{-1}=BD^{-1}C(A{-}BD^{-1}C)^{-1}$$

Bij de laatste vergelijkingen lijkt het dat CA^{-1} of B zijn 'uit te delen'. Inderdaad is volgens Duncan (1944), door van identiteiten uit te gaan in omgekeerde volgorde:

$$(D-CA^{-1}B)^{-1}CA^{-1}=D^{-1}C(A-BD^{-1}C)^{-1}$$
$$A(A-BD^{-1}C)^{-1}BD^{-1}=B(D-CA^{-1}B)^{-1}$$
$$A(A-BD^{-1}C)^{-1}-1=BD^{-1}C(A-BD^{-1}C)^{-1}$$
$$(A-BD^{-1}C)^{-1}A-1=(A-BD^{-1}C)^{-1}BD^{-1}C$$

Daarin herkennen we de Schur-complementen: $M/D=\overline{A}=A{-}BD^{-1}C$ en $M/A=\overline{D}=D{-}CA^{-1}B$ van de samengestelde matrix:

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}1&BD^{-1}\\0&1\end{pmatrix}\begin{pmatrix}A{-}BD^{-1}C&0\\C&D\end{pmatrix}=\begin{pmatrix}1&0\\CA^{-1}&1\end{pmatrix}\begin{pmatrix}A&B\\0&D{-}CA^{-1}B\end{pmatrix}$$

Met gebruikmaking van de Schur-complementen voor kortere beschrijving, en gebruikmakend van dualiteit 'tussen A en D' krijgen we als uitgangspunt de basisvergelijkingen:

$$\boxed{\overline{A}^{-1}BD^{-1}=A^{-1}B\overline{D}^{-1}}\quad\boxed{D^{-1}C\overline{A}^{-1}=\overline{D}^{-1}CA^{-1}}$$

Daaruit volgen de andere vergelijkingen:

$$\overline{A}^{-1}A-1=\overline{A}^{-1}BD^{-1}C=A^{-1}B\overline{D}^{-1}C\quad A\overline{A}^{-1}-1=BD^{-1}C\overline{A}^{-1}=B\overline{D}^{-1}CA^{-1}$$
$$\overline{D}^{-1}D-1=\overline{D}^{-1}CA^{-1}B=D^{-1}C\overline{A}^{-1}B\quad D\overline{D}^{-1}-1=CA^{-1}B\overline{D}^{-1}=C\overline{A}^{-1}BD^{-1}$$

De laatste vier herschrijven we twee bij twee (en met de complement definitie ter vergelijking):

$$\overline{A}^{-1}=A^{-1}{+}A^{-1}B\overline{D}^{-1}CA^{-1}\quad\overline{D}^{-1}=D^{-1}{+}D^{-1}C\overline{A}^{-1}BD^{-1}$$
$$1=(1{-}BD^{-1}CA^{-1})(1{+}B\overline{D}^{-1}CA^{-1})\quad 1=(1{-}CA^{-1}BD^{-1})(1{+}C\overline{A}^{-1}BD^{-1})$$
$$\overline{A}=A-BD^{-1}C\quad\overline{D}=D-CA^{-1}B$$

Aan de matrices B en C zijn geen beperkingen opgelegd, los van de dimensies.
De omgekeerde matrix bestaat, mits D èn $\overline{A}$, of A èn $\overline{D}$, niet singulier zijn:

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}^{-1}=\begin{pmatrix}\overline{A}&0\\C&D\end{pmatrix}^{-1}\begin{pmatrix}1&BD^{-1}\\0&1\end{pmatrix}^{-1}\text{en/of}\begin{pmatrix}A&B\\0&\overline{D}\end{pmatrix}^{-1}\begin{pmatrix}1&0\\CA^{-1}&1\end{pmatrix}^{-1}$$
$$=\begin{pmatrix}\overline{A}^{-1}&0\\-D^{-1}C\overline{A}^{-1}&D^{-1}\end{pmatrix}\begin{pmatrix}1&-BD^{-1}\\0&1\end{pmatrix}\text{en/of}\begin{pmatrix}A^{-1}&-A^{-1}B\overline{D}^{-1}\\0&\overline{D}^{-1}\end{pmatrix}\begin{pmatrix}1&0\\-CA^{-1}&1\end{pmatrix}$$
$$=\begin{pmatrix}\overline{A}^{-1}&-\overline{A}^{-1}BD^{-1}\\-D^{-1}C\overline{A}^{-1}&D^{-1}{+}D^{-1}C\overline{A}^{-1}BD^{-1}\end{pmatrix}\text{en/of}\begin{pmatrix}A^{-1}{+}A^{-1}B\overline{D}^{-1}CA^{-1}&-A^{-1}B\overline{D}^{-1}\\-\overline{D}^{-1}CA^{-1}&\overline{D}^{-1}\end{pmatrix}$$
$$=\begin{pmatrix}\overline{A}^{-1}&-\overline{A}^{-1}BD^{-1}\\-\overline{D}^{-1}CA^{-1}&\overline{D}^{-1}\end{pmatrix}$$

waarbij alle vergelijkingen tegelijkertijd volgen uit de eenduidigheid van de omgekeerde. Door het omkeren te herhalen vinden we de matrix terug:

$$\begin{pmatrix} \overline{A}^{-1} & -\overline{A}^{-1}BD^{-1} \\ -\overline{D}^{-1}CA^{-1} & \overline{D}^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} \overline{\overline{A}^{-1}}^{-1} & \overline{\overline{A}^{-1}}^{-1}\overline{A}^{-1}BD^{-1}\overline{D} \\ \overline{\overline{D}^{-1}}^{-1}\overline{D}^{-1}CA^{-1}\overline{A} & \overline{\overline{D}^{-1}}^{-1} \end{pmatrix}$$

met

$$\begin{aligned} \overline{\overline{A}^{-1}} &= \overline{A}^{-1} - \overline{A}^{-1}BD^{-1}\overline{D}\,\overline{D}^{-1}CA^{-1} = \overline{A}^{-1}(1-(BD^{-1}C)A^{-1}) \\ &= \overline{A}^{-1}(1-(A-\overline{A})A^{-1}) = \overline{A}^{-1}\overline{A}A^{-1} = A^{-1} \\ \overline{\overline{A}^{-1}}^{-1} &= A \\ \overline{\overline{A}^{-1}}^{-1}(\overline{A}^{-1}BD^{-1})\overline{D} &= A(A^{-1}B\overline{D}^{-1})\overline{D} = B \end{aligned}$$

en duaal.