formula for sample variance

$$\sigma_x^2 = \frac{1}{N^2}\left(n\sum_i x_i^2 + (\Sigma x_i)^2\right)$$

$$= \frac{1}{N}\left(\sum_i (x_i - \bar{x})^2\right) =$$

$$= \frac{1}{N}\left(\Sigma (x_i^2 + \bar{x}^2 - 2x_i\bar{x})\right)$$

$$= \frac{1}{N}\left(\Sigma x_i^2 + \bar{x}^2 \sum_i 1 - 2\bar{x}\Sigma x_i\right)$$

$$= \frac{1}{N}\left(\Sigma x_i^2 + N\bar{x}^2 - 2N\bar{x}^2\right)$$

$$= \frac{1}{N}\left(\Sigma x_i^2 - N\bar{x}^2\right)$$

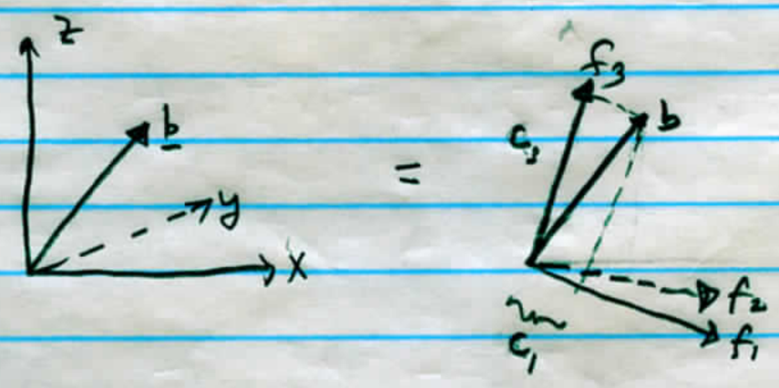
$$= \frac{1}{N}\left(\Sigma x_i^2 - \frac{1}{N}(\Sigma x_i)^2\right)$$

$$= \frac{1}{N^2}\left(N\Sigma x_i^2 - (\Sigma x_i)^2\right)$$

Transformations

$$\begin{array}{c} \\ [-\underline{s}-] = [-c-] \end{array} \begin{array}{c} F \\ \begin{bmatrix} [-f_1-] \\ [-f_2-] \\ [-f_3-] \end{bmatrix} \end{array}$$

$$[-\underline{s}-] = c_1 [-f_1-] + c_2 [-f_2-] + c_3 [-f_3-]$$



interpretation. If all f's perpendicular, then $\underline{f}$'s a new, rotated set of axes and c's are components of $\underline{s}'$ in the new, rotated coordinate system.

$$\underline{s} = \underline{c}\, \underline{\underline{F}}$$

$$\underline{c} = \underline{s}\, \underline{\underline{F}}^{-1}$$

(unary)

but if $|\underline{f}^{(i)}| = 1$ and $\underline{f}^{(i)} \cdot \underline{f}^{(j)} = 0$ then $\underline{\underline{F}}^{-1} = \underline{\underline{F}}^{T}$

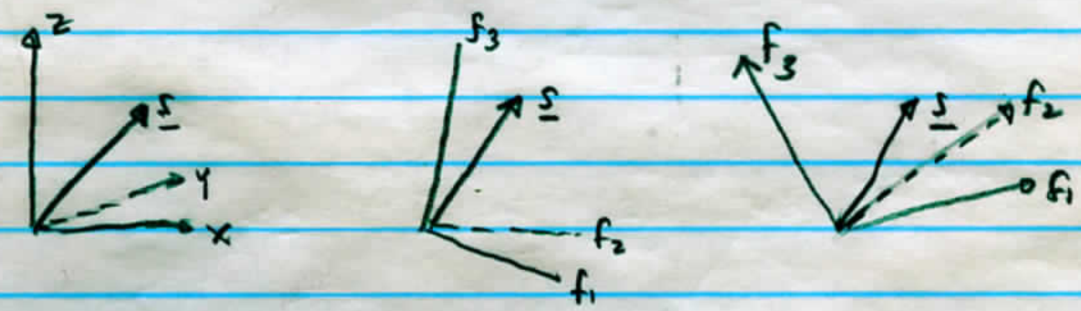
$$\begin{bmatrix} [-f_1-] \\ [-f_2-] \\ [-f_3-] \end{bmatrix} \begin{bmatrix} \begin{bmatrix} f_1 \end{bmatrix} & \begin{bmatrix} f_2 \end{bmatrix} & \begin{bmatrix} f_3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} f_1 \cdot f_1 & f_1 \cdot f_2 & f_1 \cdot f_3 \\ f_2 \cdot f_1 & f_2 \cdot f_2 & \cdots \\ & \cdots & \end{bmatrix} = I$$

$\underline{s} = \underline{c}\,\underline{\underline{F}}$ $\quad \underline{c} = \underline{s}\,\underline{\underline{F}}^T$ where $\underline{\underline{F}}$ is unary

suppose you have another unary matrix R, then

$$\underline{s} = \underline{c}\,\underline{\underline{F}} = \underline{c}\,\underline{\underline{I}}\,\underline{F} = (c R^T)(R F) = c' F'$$

F' = rotated factors



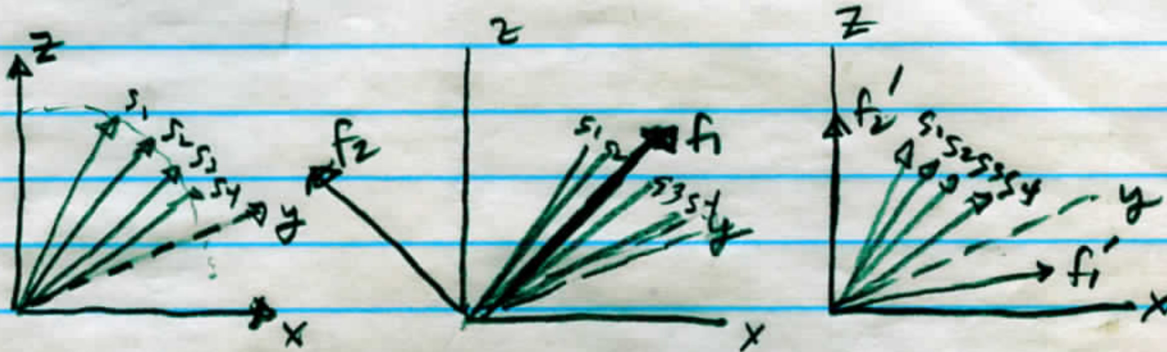
Orthogonal Rotation about 2 axes
(Given's Transformation)

$$\overset{F'}{\begin{bmatrix} [\,-\, f_1 \,-\,] \\ \cos\theta f_2 + \sin\theta f_3 \\ -\sin\theta f_2 + \cos\theta f_3 \\ [\,-\, f_4 \,-\,] \\ [\,-\, f_5 \,-\,] \end{bmatrix}} = R \overset{F}{\begin{bmatrix} [\,-\, f_1 \,-\,] \\ [\,-\, f_2 \,-\,] \\ [\,-\, f_3 \,-\,] \\ [\,-\, f_4 \,-\,] \\ [\,-\, f_5 \,-\,] \end{bmatrix}}$$

so $\underline{\underline{R}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 & 0 \\ 0 & -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Picking Good factors

SVD: orthonormal factors
in order of importance
but not



Definition of a good set of factors.

$\underline{\underline{S}}$ contains just a few ($p<m$) factors
factors near original axes?
factors mutually perpendicular?
factors unlike each other?
coefficients typically large or zero?
no factors have positive elements?
coefficients positive?

No single good choice

factors unlike one another

$$\text{minimize} \quad \sum_{s=1}^{p} \sum_{t>s}^{p} \sum_{j}^{N} f'^{2}_{js} \, f'^{2}_{jt} = M_1$$

sum over distinct pairs of factors

sum over elements

product small when atleast one factor is small

just a few big factors

$$\text{maximize} \quad \sum_{s=1}^{p} \sum_{j=1}^{N} f'^{4}_{js} = M_2$$

for orthogonal rotations, these conditions the same

$$\left(\sum_{j} f^2_{sj}\right)^2 = \sum_{j} f^4_{sj} + 2 \sum_{s=1}^{p} \sum_{t>s}^{p} f^2_{sj} f^2_{tj}$$

eg $(f^2_{s1} + f^2_{s2} + f^2_{s3})^2 = f^4_{s1} + f^4_{s2} + f^4_{s3} + 2(f_{s1}f_{s2} + f_{s1}f_{s3} + f_{s2}f_{s3})$

since l.h.s. is just 4-th power of length of f, it is unchanged under rotation. So minimizing M_1 same as maximizing M_2.

Varimax Orthogonal Rotation

find the rotation R that
maximizes the variance of squared elements of factors

$$\max\ V_s = \frac{1}{n^2}\left[n \sum f'^{4}_{sj} - \left(\sum f'^{2}_{sj}\right)^2\right]$$

for every pair of factors $\underline{f}_s$ and $\underline{f}_t$

process: sequence of givens rotations

$$\begin{array}{c} f' \\ \begin{bmatrix} \cos\theta f_1 + \sin\theta f_2 \\ -\sin\theta f_1 + \cos\theta f_2 \\ [-f_3-] \\ [-f_4-] \\ [-f_5-] \end{bmatrix} \end{array} \quad \theta_1 \quad \begin{bmatrix} [-f_1-] \\ [-f_2-] \\ [-f_3-] \\ [-f_4-] \\ [-f_5-] \end{bmatrix}$$

$$\begin{bmatrix} \cos\theta f'_1 + \sin\theta f'_3 \\ -f'_2- \\ -\sin\theta f'_1 + \cos\theta f'_3 \\ -f_4- \\ -f_5- \end{bmatrix} \quad \theta_2 \quad \begin{bmatrix} -f'_1- \\ -f'_2- \\ -f_3- \\ -f_4- \\ -f_5- \end{bmatrix}$$

etc.

each time pick θ to maximize V_1.
After completing $\underline{f}_1$, repeat for f_2 (but only mix factors $f_3, f_4 \cdots$ into f_2, $f_4, f_5 \cdots$ into f_3, etc.

The formula for θ

$$\max: V_{st} = n^2(V_s + V_t) = n\sum_j^m f_{sj}'^4 - \left(\sum_j^m f_{sj}'^2\right)^2 + n\sum_j^m f_{tj}'^4 - \left(\sum_j^m f_{tj}'^2\right)^2$$

where $f_s' = \cos\theta\, f_s + \sin\theta\, f_t$
$f_t' = -\sin\theta\, f_s + \cos\theta\, f_t$

$\frac{\partial V_{st}}{\partial \theta} = 0$ is condition for maximum

gives (after much algebra)

$$\theta = \frac{1}{4}\tan^{-1}\frac{2n\sum u_i v_i - \sum u_i \sum v_i}{n\sum_i (u_i^2 - v_i^2) - [(\sum u_i)^2 - (\sum v_i)^2]}$$

$u_i = f_{si}^2 + f_{ti}^2$ No - shold be minus

$v_i = 2 f_{si} f_{ti}$

QA 278.5

QA
278.5 Reyment & Jöreskog App. Fac. An. p 340

$u = x_j x_j - y_j y_j$

$v = 2 x_j y_j$

$A = \Sigma u_j$

$B = \Sigma v_j$

$C = (\Sigma u_j^2 - \Sigma v_j^2)$

$D = 2 \Sigma u_j v_j$

$num = D - 2AB/n$

$den = C - (A^2 - B^2)/n$

$\tan 4p = num/den$