Estimation of L and Ψ

- Based on the sample covariance matrix **S** or sample correlation matrix **R** from observations $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.
- Two most popular methods: the **principal component method** and the **maximum likelihood method**.
- The solution from either method can be rotated in order to simplify the interpretation of factors.
- Try more than one method to see if the solutions are consistent with one another.
- For large data sets, split them in half and perform a factor analysis on each part to check the stability of the solution.

The Principal Component Method

PRINCIPAL COMPONENT SOLUTION OF THE FACTOR MODEL

The principal component factor analysis of the sample covariance matrix **S** is specified in terms of its eigenvalue–eigenvector pairs $(\hat{\lambda}_1, \hat{\mathbf{e}}_1), (\hat{\lambda}_2, \hat{\mathbf{e}}_2), \ldots, (\hat{\lambda}_p, \hat{\mathbf{e}}_p)$, where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_p$. Let $m < p$ be the number of common factors. Then the matrix of estimated factor loadings $\{\tilde{\ell}_{ij}\}$ is given by

$$\tilde{\mathbf{L}} = \left[\sqrt{\hat{\lambda}_1}\,\hat{\mathbf{e}}_1 \,\vdots\, \sqrt{\hat{\lambda}_2}\,\hat{\mathbf{e}}_2 \,\vdots\, \cdots \,\vdots\, \sqrt{\hat{\lambda}_m}\,\hat{\mathbf{e}}_m\right] \qquad (9\text{-}15)$$

The estimated specific variances are provided by the diagonal elements of the matrix $\mathbf{S} - \tilde{\mathbf{L}}\tilde{\mathbf{L}}'$, so

$$\tilde{\boldsymbol{\Psi}} = \begin{bmatrix} \tilde{\psi}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\psi}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\psi}_p \end{bmatrix} \quad \text{with} \quad \tilde{\psi}_i = s_{ii} - \sum_{j=1}^{m} \tilde{\ell}_{ij}^2 \qquad (9\text{-}16)$$

Communalities are estimated as

$$\tilde{h}_i^2 = \tilde{\ell}_{i1}^2 + \tilde{\ell}_{i2}^2 + \cdots + \tilde{\ell}_{im}^2 \qquad (9\text{-}17)$$

The principal component factor analysis of the sample correlation matrix is obtained by starting with **R** in place of **S**.

Example 9.4

- Factor analysis of stock-price data
- Use m=2 and sample correlation matrix **R**.

Example 9.4 (Cont.)

Estimated factor loadings, communalities and specific variances for one-factor (m = 1) and two-factor (m = 2) decomposition

Variable	One-factor solution: Estimated factor loadings F_1	One-factor solution: Specific variances $\tilde{\psi}_i = 1 - \tilde{h}_i^2$	Two-factor solution: Estimated factor loadings F_1	Two-factor solution: Estimated factor loadings F_2	Two-factor solution: Specific variances $\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. Allied Chemical	.783	.39	.783	−.217	.34
2. Du Pont	.773	.40	.773	−.458	.19
3. Union Carbide	.794	.37	.794	−.234	.31
4. Exxon	.713	.49	.713	.472	.27
5. Texaco	.712	.49	.712	.524	.22
Cumulative proportion of total (standardized) sample variance explained	.571		.571	.733	

Market factor

Industry factor

Example 9.3 : Factor Analysis for Consumer-Preference Data

In a consumer-preference study, a random sample of customers were asked to rate several attributes of a new product. The responses, on a 7 semantic differential scale, were tabulated and the attribute correlation matrix constructed.

Attribute (Variable)		1	2	3	4	5
Taste	1	1.00	.02	.96	.42	.01
Good buy for money	2	.02	1.00	.13	.71	.85
Flavor	3	.96	.13	1.00	.50	.11
Suitable for snack	4	.42	.71	.50	1.00	.79
Provides lots of energy	5	.01	.85	.11	.79	1.00

- Var 1 and var 3, var 2 and var 5 form groups, i.e. are highly correlated.
- Var 4 is "closer" , i.e. more correlated, to (var 2, var 5).
- We might expected there are two or three common factors.

The first two eigenvalues, $\lambda_1 = 2.85$ and $\lambda_2 = 1.81$, of $\mathbf{R}$ are the only eigenvalues greater than unity. Moreover, $m = 2$ common factors will account for a cumulative proportion

$$\frac{\hat{\lambda}_1 + \hat{\lambda}_2}{p} = \frac{2.85 + 1.81}{5} = .93$$

Example 9.3 (Cont.)

Variable	Estimated factor loadings $\tilde{\ell}_{ij} = \sqrt{\hat{\lambda}_i}\,\hat{e}_{ij}$ F_1	F_2	Communalities $\tilde{h}_i^2$	Specific variances $\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. Taste	.56	.82	.98	.02
2. Good buy for money	.78	−.53	.88	.12
3. Flavor	.65	.75	.98	.02
4. Suitable for snack	.94	−.11	.89	.11
5. Provides lots of energy	.80	−.54	.93	.07
Eigenvalues	2.85	1.81		
Cumulative proportion of total (standardized) sample variance	.571	.932		

- A rotation of the factors often reveals a simple structure and aids interpretation.
- We shall re-consider this example again after factor rotation.

Maximum Likelihood Method

- Assume the common factors **F** and the errors **ε** are normally distributed.
- The MLE of **L** and **Ψ** are obtained by numerically maximizing the likelihood function subject to that $\mathbf{L}^T\mathbf{\Psi}^{-1}\mathbf{L}$ is a diagonal matrix. The constraint is a computationally convenient uniqueness condition.
- Example 9.5: factor analysis of stock-price data using the maximum likelihood method (use correlation matrix **R**)

91

Factor Rotation

- Since the originally estimated loadings may not be readily interpretable, it is usual practice to rotate them until a “simple structure” is achieved.
- The general principal is to see a pattern of loadings such that each variable loads highly on a single factor and has small to moderate loadings on the remaining factors.
- Geometric interpretation for m=2
- Varimax criterion (details not required)
- Oblique rotations (not required)

92

Example 9.10 : Factor Rotation for Stock-Price Data

Variable	Maximum likelihood estimates of factor loadings F_1	F_2	Rotated estimated factor loadings F_1^*	F_2^*	Specific variances $\hat{\psi}_i = 1 - \hat{h}_i^2$
Allied Chemical	.684	.189	.601	.377	.50
Du Pont	.694	.517	.850	.164	.25
Union Carbide	.681	.248	.643	.335	.47
Exxon	.621	−.073	.365	.507	.61
Texaco	.792	−.442	.208	.883	.18
Cumulative proportion of total sample variance explained	.485	.598	.335	.598	

Example 9.9 : Factor Rotation for Consumer-Preference Data

It is clear that variables 2, 4, and 5 define factor 1 (high loadings on factor 1, small or negligible loadings on factor 2), while variables 1 and 3 define factor 2 (high loadings on factor 2, small or negligible loadings on factor 1). Variable 4 is most closely aligned with factor 1, although it has aspects of the trait represented by factor 2. We might call factor 1 a *nutritional* factor and factor 2 a *taste* factor.

Variable	Estimated factor loadings F_1	F_2	Rotated estimated factor loadings F_1^*	F_2^*	Communalities $\hat{h}_i^2$
1. Taste	.56	.82	.02	.99	.98
2. Good buy for money	.78	−.52	.94	−.01	.88
3. Flavor	.65	.75	.13	.98	.98
4. Suitable for snack	.94	−.10	.84	.43	.89
5. Provides lots of energy	.80	−.54	.97	−.02	.93
Cumulative proportion of total (standardized) sample variance explained	.571	.932	.507	.932	

Example 9.9 (Cont.)

The factor loadings for the variables are pictured with respect to the original and (varimax) rotated factor axes in Figure 9.2. ■

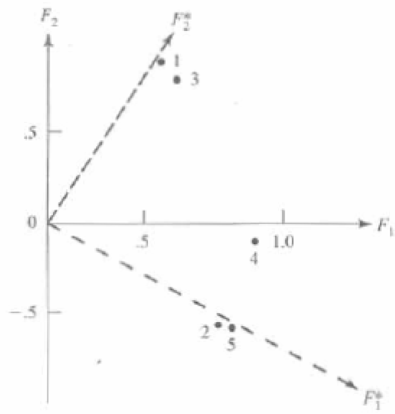
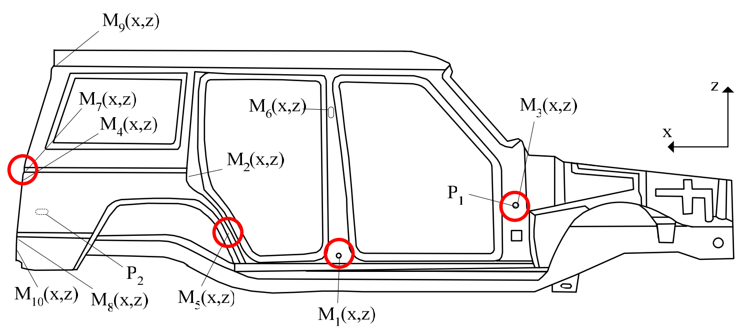


Figure 9.2 Factor rotation for hypothetical marketing data.

Example: Factor Analysis for Single Station Assembly Process Diagnosis

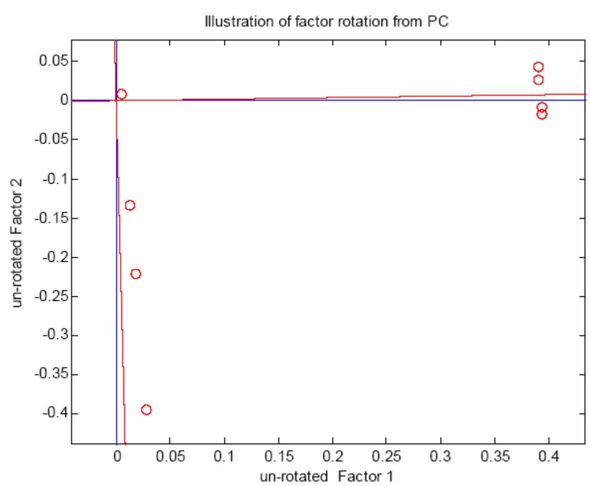


Observable variables (**x**) are 8 dimensional measurements on the assembly, including x- and z- direction deviations of measurement points M_1, M_3, M_5, and M_7. Two (unknown) process faults are present: x-direction variation of locator P_1 and z-direction variation of locator P_2.

Estimated Factor Loadings for m = 2 (PC Method)

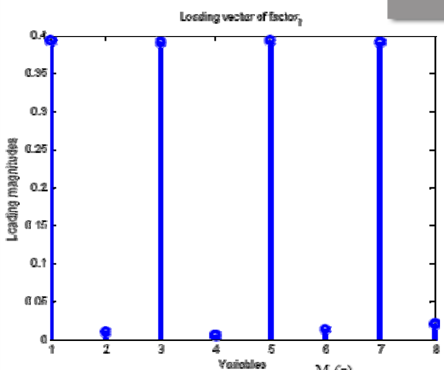
Variables	Factor 1	Factor 2	Factor 1*	Factor 2*
1	0.3934	-0.0255	0.3938	-0.0181
2	0.01	-0.1341	0.0125	-0.1339
3	0.3916	0.0187	0.3912	0.0261
4	0.0055	0.0087	0.0053	0.0088
5	0.3935	-0.0166	0.3938	-0.0091
6	0.0133	-0.2214	0.0175	-0.2211
7	0.3912	0.0361	0.3904	0.0435
8	0.0202	-0.3965	0.0277	-0.396

Factor Rotation for Assembly Process Diagnosis

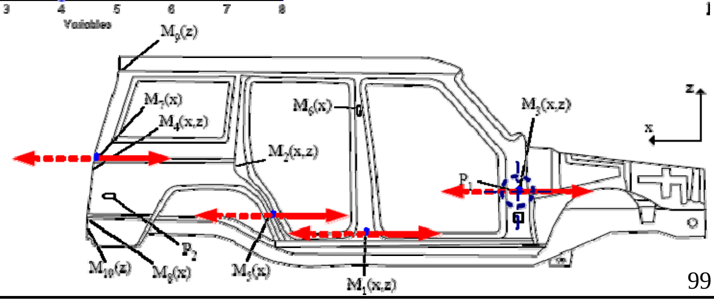


$$\mathbf{T} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} 0.9998 & -0.0189 \\ 0.0189 & 0.9998 \end{bmatrix}, \text{ and } \varphi = 1.146^\circ$$

Visualization of Loading Vectors for Factor 1



- All measurement points "move" horizontally
- With almost equal magnitudes
- Factor 1 corresponds to locator P_1's variation along x-direction.



Visualization of Loading Vectors for Factor 2

- All measurement points, except M_3, "move" along circles centered at P_1.
- With magnitudes proportional to the distance from it to P_1
- Factor 2 corresponds to locator P_2's variation along z-direction.