

マーケティング・リサーチ特論 ～因子分析～

2024年度1学期： 水曜3限
担当教員： 石垣 司

因子分析

多変量データを発生する要因となっている隠れた因子を見つけるための手法(Spearman 1904)

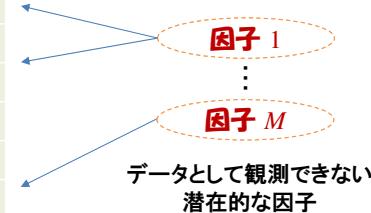
- 変数間の相関構造から潜在因子を推定
- 心理学, 社会学, マーケティングで多用



Charles Spearman 1863-1945

多変量データ

| ID | 変数1 | 変数2 | ... | 変数 P |
|-----|----------|----------|-----|----------|
| 1 | x_{11} | x_{12} | ... | x_{1P} |
| 2 | x_{21} | x_{22} | ... | x_{2P} |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| i | x_{i1} | x_{i2} | ... | x_{iP} |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| N | x_{N1} | x_{N2} | ... | x_{NP} |

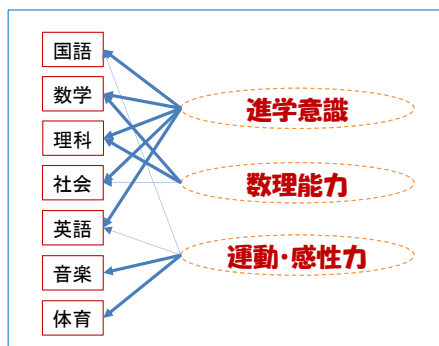


因子分析のイメージ

潜在的な因子の探索や実証

- 例: 300人分の中学校のテスト

| 生徒 No. | 国語 | 数学 | 理科 | 社会 | 英語 | 音楽 | 体育 |
|--------|----|----|----|----|----|----|----|
| 1 | 83 | 60 | 55 | 81 | 90 | 50 | 93 |
| 2 | 70 | 80 | 78 | 80 | 55 | 44 | 59 |
| 3 | 50 | 90 | 95 | 70 | 80 | 80 | 49 |
| 4 | 60 | 44 | 44 | 99 | 78 | 73 | 30 |
| 5 | 57 | 80 | 80 | 50 | 67 | 64 | 59 |
| 6 | 55 | 65 | 70 | 65 | 67 | 30 | 70 |
| 7 | 80 | 73 | 66 | 46 | 55 | 58 | 88 |
| 8 | 98 | 40 | 50 | 88 | 99 | 93 | 54 |
| 9 | 55 | 77 | 88 | 40 | 89 | 88 | 97 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |



直接観測できない潜在的な因子(構成概念)を探る

- 因子の探索: データから因子を発見
- 理論の実証: データが理論に合致するかの検証

因子分析による知覚マップ

例: ビールのイメージ調査

岡田、守口「マーケティングのデータ分析」朝倉書店(2010)の分析用データ(529人)を使用した分析結果

5段階尺度(0~4点) 529人の合計得点

$$p = 1, \dots, P$$

| ブランド名 | ノド越し | 香り | 味 | 幸せ | ホッ | 爽快 |
|-------|------|------|------|------|------|------|
| AKライ | 1628 | 1305 | 1435 | 1335 | 1275 | 1533 |
| Kラガー | 1439 | 1428 | 1443 | 1339 | 1342 | 1366 |
| Sエビス | 1508 | 1631 | 1638 | 1534 | 1486 | 1376 |
| K淡麗 | 1443 | 1213 | 1261 | 1226 | 1243 | 1337 |
| 金麦 | 1224 | 1207 | 1197 | 1191 | 1219 | 1181 |
| K一番 | 1571 | 1548 | 1589 | 1460 | 1456 | 1482 |
| S黒ラベル | 1419 | 1418 | 1417 | 1349 | 1354 | 1356 |
| フレモル | 1562 | 1599 | 1608 | 1544 | 1512 | 1493 |
| A本生 | 1334 | 1195 | 1219 | 1220 | 1229 | 1294 |
| Kのどし | 1385 | 1207 | 1209 | 1216 | 1235 | 1309 |

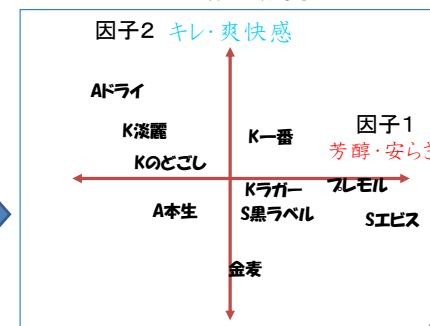
$i = 1, \dots, N$

$$j = 1, \dots, M$$

| 因子負荷量行列(プロマックス回転) | 因子1 | 因子2 |
|-------------------|-------|-------|
| ノド越し | 0.000 | 0.981 |
| 香り | 1.048 | 0.000 |
| 味 | 0.848 | 0.205 |
| 幸せ | 0.921 | 0.102 |
| ホッとする | 1.027 | 0.000 |
| 爽快 | 0.000 | 0.987 |

知覚マップによるポジショニング

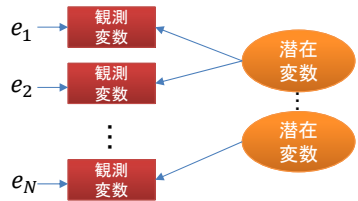
- 因子の解釈
 - 因子1: “芳醇・安らぎ”軸
 - 因子2: “キレ・爽快感”軸
- ポジショニング
 - 市場内でのブランドの位置づけ
 - ブランドの“顔”の明確化



因子分析と主成分分析の違い

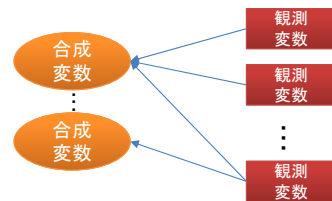
因子分析

- 潜在変数の推定



主成分分析

- 変数の合成



- 両者の発想は真逆のため使い分けが必要
 しかしながら、同じ多変量データを分析した場合、得られる結果は同傾向になることが多い
 使い分けの許容度は利用する分野により様々

因子分析モデル

数理的な表記法 ※データではなく確率変数としての表記であることに注意

- 標準化した観測変数: x_p ($p = 1, \dots, P$)
 各変数毎(“ノド越し”や“香り”など)に値を標準化(平均0, 分散1)
- 共通因子(因子スコア)(確率変数): f_m ($m = 1, \dots, M$)
- 因子負荷量(推定すべき係数パラメータ): a_{pm}
- 独自因子(確率変数): e_p (誤差項)
- 例: ビールイメージ調査

6 変数($P = 6$)から 2 因子を抽出するモデル

$$\begin{matrix} x_1: \text{ノド越し} \\ x_2: \text{香り} \\ x_3: \text{味} \\ x_4: \text{幸せ} \\ x_5: \text{ホッとする} \\ x_6: \text{爽快} \end{matrix} \begin{cases} x_1 = a_{11}f_1 + a_{12}f_2 + e_1 \\ x_2 = a_{21}f_1 + a_{22}f_2 + e_2 \\ \vdots \\ x_6 = a_{61}f_1 + a_{62}f_2 + e_6 \end{cases}$$

因子分析モデル(行列表記)

- 標準化した観測変数ベクトル: $x = [x_1, \dots, x_p]^T$
- 共通因子ベクトル: $f = [f_1, \dots, f_M]^T$
- 因子負荷量行列: $A = \begin{bmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{p1} & \dots & a_{pM} \end{bmatrix}$
- 誤差ベクトル: $e = [e_1, \dots, e_p]^T$

因子分析モデル: $x = Af + e$

観測データ $\{x_i\}_{i=1, \dots, N}$ からパラメータ行列 A を推定

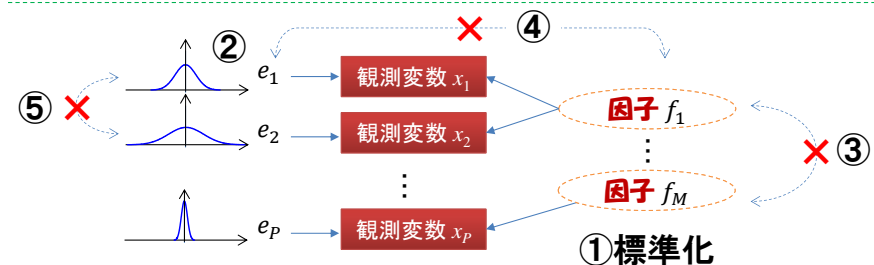
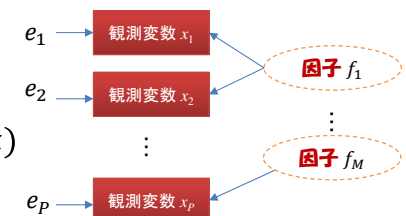
- パラメータ A も確率変数 f も未知のため、モデル構造に制約(仮定)を加えてパラメータを推定できるようにする

因子分析モデルの仮定

モデルの仮定

- ① $E[f_m] = 0, V[f_m] = 1$
- ② $e_p \sim N(0, \sigma_p^2)$
- ③ $\text{Cov}[f_m, f_k] = 0$ ($m \neq k$)
- ④ $\text{Cov}[f_m, e_p] = 0$
- ⑤ $\text{Cov}[e_p, e_q] = 0$ ($p \neq q$)

各主体 i に対して



因子負荷量(パラメータ行列A)の推定 #1

基本的方針: 母相関係数と標本相関を近づける

- 手段: 主因子法, 最小2乗法, 最尤法など

- 母相関係数行列 R (x の分散共分散行列)

(x は標準化(平均ゼロ, 分散1)されているので $R = V[x]$)

$$R = \begin{bmatrix} 1 & r_{21} & \dots & r_{p1} \\ r_{21} & 1 & & \vdots \\ \vdots & & \ddots & \\ r_{p1} & \dots & & 1 \end{bmatrix} = AA^T + \Psi \quad \Psi = E[ee^T] = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_p^2 \end{bmatrix}$$

- 観測データの標本相関係数行列 S

$$S = \begin{bmatrix} 1 & s_{21} & \dots & s_{p1} \\ s_{21} & 1 & & \vdots \\ \vdots & & \ddots & \\ s_{p1} & \dots & & 1 \end{bmatrix} =$$

| | ノド越し | 香り | 味 | 幸せ | ホッ | 爽快 |
|------|-------|-------|-------|-------|-------|-------|
| ノド越し | 1.000 | | | | | |
| 香り | 0.621 | 1.000 | | | | |
| 味 | 0.772 | 0.972 | 1.000 | | | |
| 幸せ | 0.718 | 0.979 | 0.981 | 1.000 | | |
| ホッ | 0.626 | 0.987 | 0.956 | 0.983 | 1.000 | |
| 爽快 | 0.978 | 0.589 | 0.740 | 0.691 | 0.691 | 1.000 |

9

因子負荷量(パラメータA)の推定 #2

$R = AA^T + \Psi$ の導出

$$\begin{aligned} R &= V[x] = E[xx^T] \\ &= E[(Af + e)(Af + e)^T] = E[(Af + e)(f^T A^T + e^T)] \\ &= AE[ff^T]A^T + AE[fe^T] + E[ef^T] + E[ee^T] \\ &= AE[ff^T]A^T + E[ee^T] = AA^T + \Psi \end{aligned}$$

因子負荷量行列 A の直交回転の不定性

$g = Bf$ となる正則行列 B , $C = AB^{-1}$ となる行列 C を考える
 $Af = AB^{-1}Bf \Rightarrow Af = Cg$ となり A と f の組み合わせは無数に存在

- 不定性を取り除くための制約条件

$$A^T A - \text{diag}(A^T A) = O \quad (O \text{ はゼロ行列})$$

$\text{rank}(A) = M$ ならばこの制約条件の下で A は一意に推定できる
 A の推定後に因子の回転により, より良い因子の解釈を得る

10

因子負荷量が推定できる必要条件 #1

例: 5 変数 2 因子のモデル $A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} & a_{51} \\ a_{12} & a_{22} & a_{32} & a_{42} & a_{52} \end{bmatrix}$

$$R = AA^T + \Psi = \begin{bmatrix} a_{11}^2 + a_{12}^2 & & & & \\ a_{11}a_{21} + a_{12}a_{22} & a_{21}^2 + a_{22}^2 & & & \\ \vdots & \vdots & \ddots & & \\ a_{11}a_{51} + a_{12}a_{52} & a_{21}a_{51} + a_{22}a_{52} & \dots & a_{51}^2 + a_{52}^2 & \end{bmatrix} + \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_5^2 \end{bmatrix}$$

- モデルのパラメータ数 因子負荷量10 + 分散 5 = 15個

- このモデルと制約条件から立てられる方程式 16本

| | | |
|-----------|--|--|
| 分散 (5本) | $\begin{cases} 1 = a_{11}^2 + a_{12}^2 + \sigma_1^2 \\ \vdots \\ 1 = a_{51}^2 + a_{52}^2 + \sigma_5^2 \end{cases}$ | 制約条件 (1本) $A^T A - \text{diag}(A^T A) = \begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix} \text{とおくと}$ |
| 共分散 (10本) | $\begin{cases} s_{21} = a_{11}a_{21} + a_{12}a_{22} \\ \vdots \\ s_{54} = a_{41}a_{51} + a_{42}a_{52} \end{cases}$ | 制約条件 $a = 0$ |

11

因子負荷量が推定できる必要条件 #2

モデルのパラメータ数と

- モデルのパラメータ推定: $P \times M + P$

- モデルから立てられる方程式の数: $\frac{P(P+1)}{2}$

- 制約条件から立てられる方程式の数: $\frac{M(M-1)}{2}$

因子負荷量行列(パラメータ A)を推定できる必要条件

- パラメータ数 \leq 方程式数 ($P(M+1) \leq \frac{P(P+1)+M(M-1)}{2}$)

4 変数 2 因子モデル: パラメータ数12, 方程式数10 \Rightarrow 推定不可

6 変数 3 因子モデル: パラメータ数24, 方程式数24 \Rightarrow 推定可能性あり

6 変数 4 因子モデル: パラメータ数24, 方程式数27 \Rightarrow 推定不可

#1 必要条件であり, この条件を満たすと必ず識別可能であるとは言っていない
 #2 その他の必要条件や十分条件も知られているが複雑なため本授業ではこれ以上取り扱わない

12

最尤法による因子負荷量の推定

因子分析の最尤法

- 対数尤度関数を最大化する \hat{A} と $\hat{\Psi}$ を推定
 - x は多変量正規分布に従って発生すると仮定
 - \hat{A} と $\hat{\Psi}$ を交互に反復計算 (Newton法や準Newton法が利用可)
 - R の factanal 関数のデフォルト設定は最尤法

確率変数 $x \sim N_p(0, R)$
 x の同時密度関数 $f(x) = \frac{1}{(\sqrt{2\pi})^N \sqrt{|R|}} \exp\left(-\frac{1}{2}x^T R^{-1}x\right)$
 データ $\{x_i\}$ の対数尤度関数 $\log \prod_{i=1}^N f(x_i) = -\frac{N}{2} \{P \log(2\pi) + \log |R| + \text{tr}(R^{-1}S)\}$

- 推定結果の例:
 因子数2と設定

| 因子負荷量行列(プロマックス回転) | | |
|-------------------|-------|-------|
| | 因子1 | 因子2 |
| ノド越し | 0.000 | 0.981 |
| 香り | 1.048 | 0.000 |
| 味 | 0.848 | 0.205 |
| 幸せ | 0.921 | 0.102 |
| ホッとする | 1.027 | 0.000 |
| 爽快 | 0.000 | 0.987 |

➡ 因子の解釈

因子の回転

回転の不定性

- 行列 B の選択次第で、解の組み合わせは無数に存在

再掲: $g = Bf$ となる正則行列 B , $C = AB^{-1}$ となる行列 C を考える
 $Af = AB^{-1}Bf \Rightarrow Af = Cg$ となり A と f の組み合わせは無数に存在

因子負荷量推定時の制約条件は推定の一意性を確保するために導入しただけであり、推定の結果を解釈しやすい直交軸になっているとは限らない

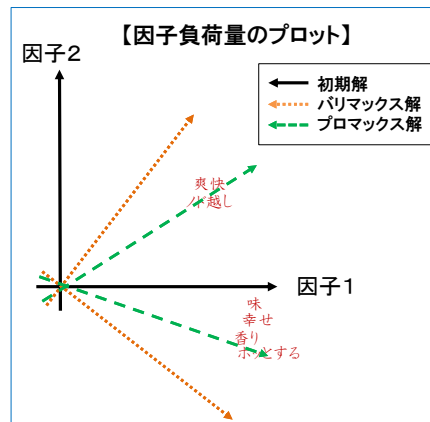
結果をより解釈しやすいように軸を回転

- 直交回転(直交軸による回転)
 - 代表例: プロマックス回転
- 斜交回転(斜交軸による回転)
 - 代表例: バリマックス回転

直交回転と斜交回転の例

推定された因子負荷量

- 初期解
 - 因子間の違いがわかりにくい
- 直交解(バリマックス回転)
 - 因子間の違いがわかりやすい
- 斜交解(プロマックス回転)
 - 直交解よりも違いが鮮明
 - ※ 因子間の相関を認めているため (仮定③を緩和している)



| 因子負荷量行列(最尤法:初期解) | | |
|------------------|-------|--------|
| | 因子1 | 因子2 |
| ノド越し | 0.848 | 0.525 |
| 香り | 0.940 | -0.335 |
| 味 | 0.987 | -0.125 |
| 幸せ | 0.971 | -0.201 |
| ホッとする | 0.936 | -0.318 |
| 爽快 | 0.820 | 0.538 |

| 因子負荷量行列(バリマックス回転) | | |
|-------------------|-------|-------|
| | 因子1 | 因子2 |
| ノド越し | 0.371 | 0.926 |
| 香り | 0.956 | 0.288 |
| 味 | 0.869 | 0.485 |
| 幸せ | 0.901 | 0.414 |
| ホッとする | 0.943 | 0.299 |
| 爽快 | 0.341 | 0.920 |

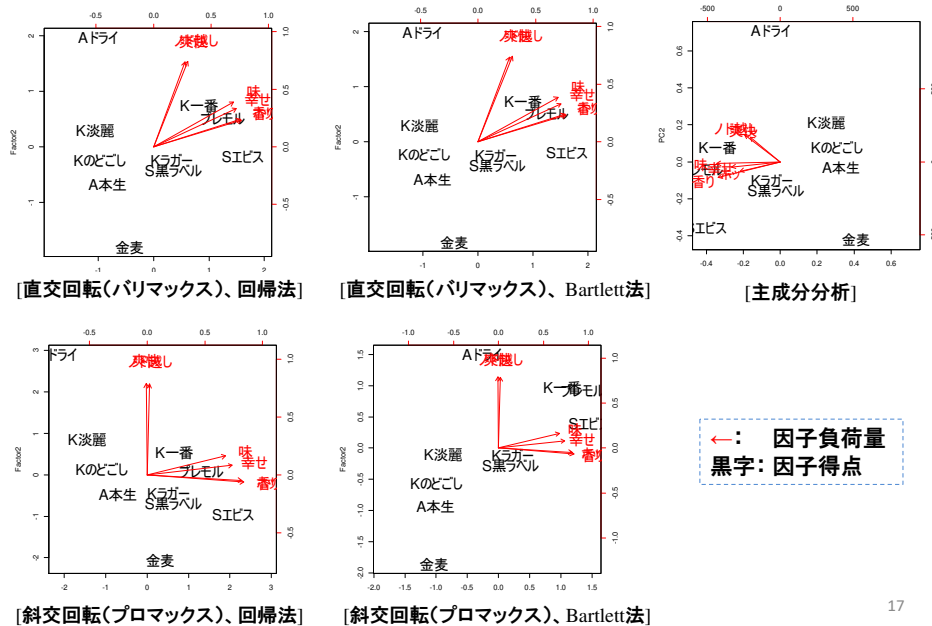
| 因子負荷量行列(プロマックス回転) | | |
|-------------------|-------|-------|
| | 因子1 | 因子2 |
| ノド越し | 0.000 | 0.981 |
| 香り | 1.048 | 0.000 |
| 味 | 0.848 | 0.205 |
| 幸せ | 0.921 | 0.102 |
| ホッとする | 1.027 | 0.000 |
| 爽快 | 0.000 | 0.987 |

主体 i の因子スコア f_i の予測

各主体 i 毎の共通因子 $\{f_{ij}\}$ の予測

- f_i は知覚マップ上の各ブランド i の座標点
 f_i は母数ではなく確率変数。推定ではなく予測
 \hat{A}, \hat{R} : 推定された因子負荷行列, 相関係数行列
 $x_i = \hat{A}f_i + e_i$ (x_i と \hat{A} は所与) から \hat{f}_i を予測
- 予測法の例1: 回帰法 Thurstone, L. L. (1935) [check!](#)
 平均2乗誤差行列 $MSE(f, \hat{f}_i) \equiv E[(f - \hat{f}_i)(f - \hat{f}_i)^T]$ の各要素を最小化
 推定量 $\hat{f}_i = \hat{A}^T \hat{R}^{-1} x_i$
- 予測法の例2: Bartlett法 Bartlett, M. S. (1937) [check!](#)
 尤度関数の最大化(独自因子 e_i の最小化)
 推定量 $\hat{f}_i = (\hat{A}^T \hat{R}^{-1} \hat{A})^{-1} \hat{A}^T \hat{R}^{-1} x_i$

推定法・回転法による結果の違い



Rでの実行例

4ページ目のデータを利用

```
#データの入力
Data = read.csv("beer.csv",header=T)
Beer = Data[,2:ncol(Data)]
rownames(Beer) = Data[,1]
```

```
#バリマックス回転+回帰法
FAvarimax = factanal(Beer, factors=2,
rotation="varimax", scores = "regression")
biplot(FAvarimax$scores,FAvarimax$loadings)
```

```
#プロマックス回転+回帰法、
FApromax = factanal(Beer, factors=2,
rotation="promax", scores = "regression")
biplot(FApramax$scores,FApramax$loadings)
```

```
#バリマックス回転+Bartlett法
FAvarimax2 = factanal(Beer, factors=2,
rotation="varimax", scores = "Bartlett")
biplot(FAvarimax2$scores,FAvarimax2$loadings)
```

```
#プロマックス回転+Bartlett法、
FApromax2 = factanal(Beer, factors=2,
rotation="promax", scores = "Bartlett")
biplot(FApramax2$scores,FApramax2$loadings)
```

因子分析の流儀

探索的因子分析

- 帰納的データ分析(本日のビールの知覚マップはこちら)
- あらかじめ因子数を決定しない分析
- なんらかの基準で因子数を決定
固有値基準, プロット法, 累積寄与率

検証的因子分析

- 理論の検証アプローチ(次回の構成概念はこちら)
- あらかじめ因子数が分かっている場合の分析
- なんらかの理論に基づき因子数が既知
心理学的知見などの各分野の知見

まとめ

因子分析は多変量データを生成する潜在的な少数の因子を見つけ出したり検証したりするために利用される

マーケティング・リサーチにおけるアンケート調査の分析と相性がよい

- 今回の授業では、消費者の潜在的な認識の可視化といわれる知覚マップの作成法を紹介

マーケティング・リサーチで主に知りたい対象は直接観測することができない購入理由や態度など
⇒次回以降、アンケート調査を用いた潜在的な構成概念の測定について

補足: 因子スコア f_i の予測 #1

予測法の例1: 回帰法の導出

- $\text{MSE}(f, \hat{f}_i)$ の各要素を最小化する $\hat{f}_i = Wx_i$ となる行列 W を求める
- $\text{MSE}(f, \hat{f}_i) \equiv E[(f - \hat{f}_i)(f - \hat{f}_i)^T]$
 $= E[(f - Wx_i)(f - Wx_i)^T]$
 $= E[ff^T] - E[fx_i^T W^T] - E[Wx_i f^T] + E[Wx_i x_i^T W^T]$
 $= I - \hat{A}^T W^T - W\hat{A} + W\hat{R}W^T$
 $= (\hat{R}W^T - \hat{A})^T \hat{R}^{-1} (\hat{R}W^T - \hat{A}) - \hat{A}^T \hat{R}^{-1} \hat{A} + I$
- \hat{R}^{-1} は半正定値行列 (定理: 相関係数行列は半正定値行列 & 半正定値行列の逆行列も半正定値行列)
- $\hat{R}W^T - \hat{A}$ がゼロ行列のとき $\text{MSE}(f, \hat{f}_i)$ の各要素は最小
 $\hat{R}W^T - \hat{A} = 0 \Leftrightarrow W^T = \hat{R}^{-1} \hat{A} \Leftrightarrow W = \hat{A}^T \hat{R}^{-1}$
 $\therefore \hat{f}_i = Wx_i = \hat{A}^T \hat{R}^{-1} x_i$

$$E[x_i f^T] = E[(\hat{A}f + e_i) f^T] = \hat{A} E[ff^T] + E[e_i f^T] = \hat{A}$$
$$E[fx_i^T W^T] = E[(x_i f^T)^T W^T] = E[(Wx_i f^T)^T] = (WE[x_i f^T])^T = (W\hat{A})^T = \hat{A}^T W^T \quad 21$$

補足: 因子スコア f_i の予測 #2

予測法の例2: Bartlett法の導出

- $x_i = \hat{A}f_i + e_i$ (x_i と \hat{A} は所与) から \hat{f}_i を予測
- x_i, \hat{A}, \hat{R} は所与なので, 確率変数は e_i
- $e_i \sim N_p(0, R)$ と仮定したときの同時密度関数
$$f(e_i) = \frac{1}{(\sqrt{2\pi})^p \sqrt{|\hat{R}|}} \exp\left(-\frac{1}{2} e_i^T \hat{R}^{-1} e_i\right)$$
- x_i の対数尤度関数
$$\log\{f(e_i)\} = -\frac{1}{2} e_i^T \hat{R}^{-1} e_i + C = -\frac{1}{2} (x_i - \hat{A}f_i)^T \hat{R}^{-1} (x_i - \hat{A}f_i) + C$$
$$\frac{\partial \log\{f(e_i)\}}{\partial f_i} = \hat{A} \hat{R}^{-1} (x_i - \hat{A}f_i) = 0$$
$$\hat{A} \hat{R}^{-1} x_i = \hat{A} \hat{R}^{-1} \hat{A} f_i$$

$$\therefore \hat{f}_i = (\hat{A} \hat{R}^{-1} \hat{A})^{-1} \hat{A} \hat{R}^{-1} x_i$$

22

補足: 因子スコア f_i の予測 #3

線形回帰分析の正規方程式ではダメ?

- $x_i = \hat{A}f_i + e_i$ は x_i と \hat{A} が所与の線形回帰モデル。なぜ, 正規方程式 $\hat{f}_i = (\hat{A}^T \hat{A})^{-1} \hat{A}^T x_i$ で予測しないか?
- 因子分析モデルは観測変数 x_p 毎に誤差項の分散は異なることを許したモデル。正規方程式を用いた予測では変数 p 毎の分散の大きさを考慮していない。回帰法も Bartlett法も予測子の中に相関係数行列が入っており, 観測変数の分散の大きさが考慮されている