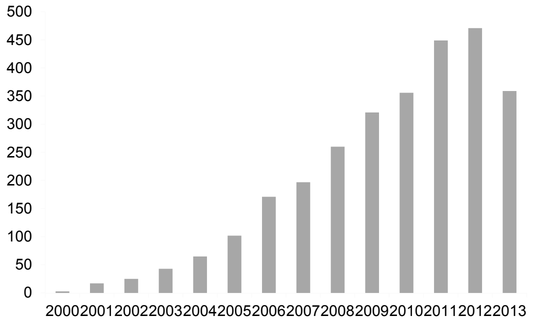
Voxel-based morphometryとは

- 特定の領域ではなく全脳を対象に、灰白質・白質の密度や容積をボクセル毎に評価する手法
- SPMを用いたVBMは、"mass univariate analysis"（多数の単変量解析）
  - 基本的な統計の理解が重要
    - 一般線形モデル（General Linear Model）
    - 多重比較補正

VBMを使った論文の推移

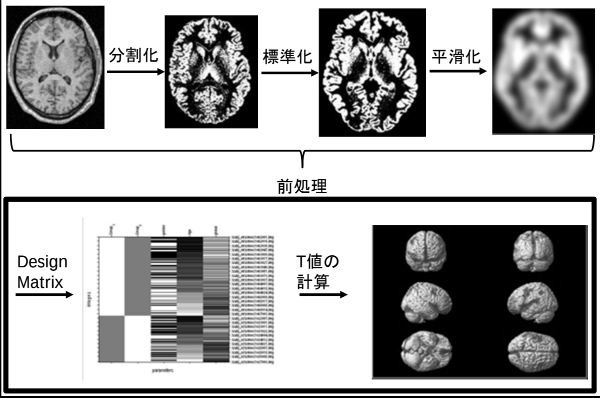


as of 14/Aug/2013

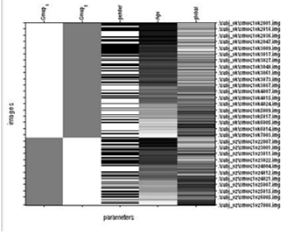
VBMでできること、できないこと

- できることの例
  - 2集団の群間比較：2標本のt検定
  - 1集団の治療前後比較：対応のあるt検定
  - 心理尺度の得点と相関する領域：相関解析
  - 遺伝子の一塩基多型の違いと疾患の有無が脳容積にもたらす影響：2要因の分散分析
- できないことの例
  - 1個人の介入前後の脳容積の変化：統計にかけられない！
  - 非常に少ない集団の群間比較：パラメトリック検定にのらない！

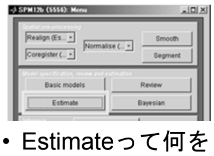
VBMの基本的な流れ



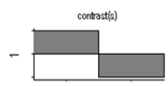
素朴な疑問



- Design matrixって何を意味するんだろう？

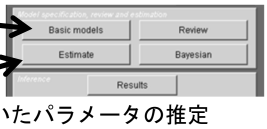


- Estimateって何をするんだろう？



- コントラストは何をしているんだろう？

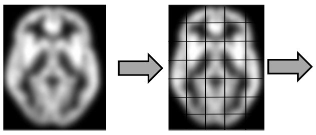
VBMにおける統計処理



- 統計モデルの作成
  - 計画行列の作成
- 推定（Estimate）
  - 一般線形モデルに基づいたパラメータの推定
- 結果表示（Results）
  - コントラストの設定
- 本講演の目的
  - SPMにおいて用いられている一般線形モデル（GLM）を通して、計画行列 (Design matrix), Estimate, コントラストについて理解する

VBMにおけるデータ処理

1. 平滑化が終わった画像からボクセル毎に数値を取り出し、行列として配置する



0	0.4	0.8	0.4	0
0.2	0.5	0.7	0.5	0.2
0.5	0.6	0.6	0.6	0.2
0.7	0.3	0.5	0.3	0.7
0.5	0.6	0.6	0.6	0.2
0.2	0.5	0.7	0.5	0.2
0	0.4	0.8	0.4	0

2. ひとつのボクセルに対して、単変量解析を行い、すべてのボクセルに対して同じ単変量解析を繰り返す(mass univariate)

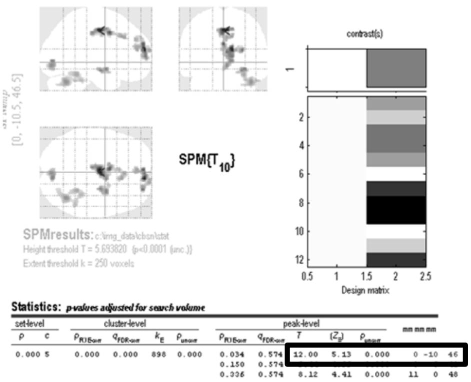
一般線形モデルとは

- 説明変数の線形結合に残差項を加えて目的変数を説明するモデル
  - 目的変数を以下の式で説明しようとする

$$\mathbf{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E}$$

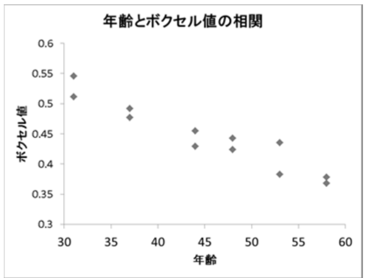
- 年齢と脳容積の相関の場合
  - Y：目的変数＝画像データ（灰白質容積）
  - X：計画行列（年齢、その他）
  - B：まだわからないパラメータ（回帰直線の傾きと切片）
  - E：回帰直線では説明しきれないゆらぎ

相関解析（年齢と負の相関を示す領域）



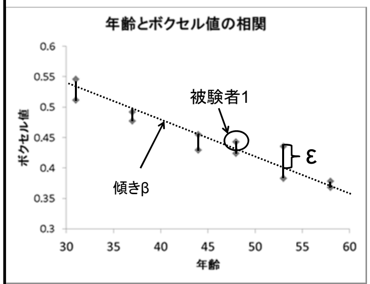
相関解析

被験者	ボクセル値 (0 -10 46)	年齢
1	0.42	48
2	0.44	53
3	0.43	44
4	0.45	44
5	0.44	48
6	0.38	58
7	0.49	37
8	0.51	31
9	0.55	31
10	0.37	58
11	0.38	53
12	0.48	37



仮説：年齢とボクセル値は負の相関を示すのではないか？

年齢とボクセル値の関係が直線で表示されると仮定

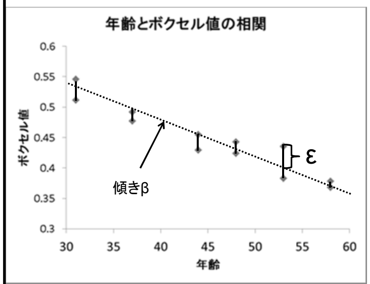


$Y_j = c + \beta \times x_j + \varepsilon_j$

- Y: ボクセル値
- j: 被験者番号
- β: 直線の傾き
- x: 年齢
- c: 切片（年齢が0の時のボクセル値）
- ε: ゆらぎ（残差）

被験者1ならば、ボクセル値0.42、年齢48歳なので、
$\mathbf{0.42} = \boldsymbol{c} + \boldsymbol{\beta} \times \mathbf{48} + \boldsymbol{\varepsilon_1}$ と表現できる。

もっともフィットする直線は？



- もっともフィットする直線は、残差εを最小にする直線
- 残差の二乗の和を求め、その値を最小にするβを見つける（最小二乗法）

$\varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2$

行列の基本

- 行列のかけ算 $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}$
- AにA の逆行列A^{-1}をかけると単位行列 I になる

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- 行列に単位行列Iをかけても何も変わらない

$$A \times I = I \times A = A$$

- AにAの転置行列A^Tをかけると正方行列になる

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 8 \\ 8 & 5 \end{bmatrix}$$

データを行列で表示

- 行列で表示することのメリット
  - 複数のボクセルに対して同じ処理を一括して行うことができる（コンピュータで扱いやすい）
  - 複雑な式を簡単な式で表すことができる
- 例：連立方程式を行列で解く

$$\begin{aligned} 2x + y &= 4 \\ 6x + 5y &= 16 \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} 2 & 1 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 16 \end{pmatrix}$$

$$AX = B$$

両辺にA^{-1}をかけるとXが簡単に求まる。

$$A^{-1} \times AX = A^{-1} \times B \qquad X = A^{-1} \times B$$

$\boldsymbol{Y_j = c + \beta \times x_j + \varepsilon_j}$ を行列で表示

$$\boldsymbol{Y_j = 1 \times c + x_j \times \beta + \varepsilon_j}$$

$$\mathrm{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E}$$

Y
0.42
0.44
0.43
0.45
0.44
0.38
0.49
0.51
0.55
0.37
0.38
0.48

データ行列

=

X	
1	48
1	53
1	44
1	44
1	48
1	58
1	37
1	31
1	31
1	58
1	53
1	37

計画行列

×

B
c
β

パラメータ行列

+

E
ε1
ε2
ε3
ε4
ε5
ε6
ε7
ε8
ε9
ε10
ε11
ε12

残差行列

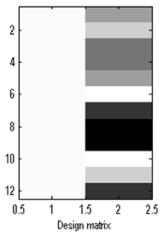
計画行列を色で表示すると

1	48
1	53
1	44
1	44
1	48
1	58
1	37
1	31
1	31
1	58
1	53
1	37

計画行列

1	48
1	53
1	44
1	44
1	48
1	58
1	37
1	31
1	31
1	58
1	53
1	37

色をつけると



SPMの画面

Eを最小にするBを求めるには

$$\mathrm{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E}$$

- もし、残差行列Eがなく、Xが正方行列であれば、BはYにXの逆行列$\boldsymbol{X}^{-1}$をかければ簡単に求まる。

$$\mathrm{Y} = \boldsymbol{X} \times \boldsymbol{B} \quad \Longrightarrow \quad \mathrm{B} = \boldsymbol{X}^{-1} \times \mathrm{Y}$$

- 残差行列Eがある場合は、最小二乗法でEを最も小さくしてBを求める方法がある。それが「疑似逆行列」を使う方法。Xの疑似逆行列$\boldsymbol{X}^{+}$は以下。

$$\boldsymbol{X}^{+} = (X^T X)^{-1} X^T$$

T: 転置行列

- このとき、推定されるBを$\widehat{\boldsymbol{B}}$とすると、$\widehat{\boldsymbol{B}}$は以下で求まる。

$$\widehat{\boldsymbol{B}} = \boldsymbol{X}^{+} \times \mathrm{Y}$$

なぜ$\boldsymbol{(X^T X)^{-1} X^T}$ なのか？

- $\widehat{\boldsymbol{B}}$を求めたい

$$\boldsymbol{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E} \quad \Rightarrow \quad \boldsymbol{Y} = \boldsymbol{X} \times \widehat{\boldsymbol{B}}$$

- $\boldsymbol{X}$ は正方行列でないため、$\boldsymbol{X}^{-1}$ は計算できない。このため、正方行列を作る工夫をする。$\boldsymbol{X}$ に転置行列$\boldsymbol{X}^{T}$をかけると、$\boldsymbol{X}^{T}\boldsymbol{X}$は正方行列になる。

$$\boldsymbol{X}^{T} \times \boldsymbol{Y} = \boldsymbol{X}^{T}\boldsymbol{X} \times \widehat{\boldsymbol{B}}$$

- $\boldsymbol{X}^{T}\boldsymbol{X}$は正方行列だから、この逆行列をかけると

$$\boldsymbol{(X^T X)^{-1} X^T} \times \boldsymbol{Y} = \boldsymbol{(X^T X)^{-1} X^T X} \times \widehat{\boldsymbol{B}}$$

$$\boldsymbol{(X^T X)^{-1} X^T} \times \boldsymbol{Y} = \widehat{\boldsymbol{B}}$$

$X^{+} = \left(X^{T}X\right)^{-1}X^{T}$にびびらない (;ﾟдﾟ)

Y	X		E		B
	定数	x	ε		切片c
4.2	1	2	ε1		傾きβ
4.8	1	3	ε2		
6.1	1	4	ε3		
(y=x+2, つまりβ=1, c=2を想定)					

$$\mathbf{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E}$$

$$\widehat{\boldsymbol{B}} = \boldsymbol{X}^{+} \times \mathbf{Y}$$

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \quad X^{T} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix}$$

$$(X^{T}X) = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 9 \\ 9 & 29 \end{bmatrix} \quad (X^{T}X)^{-1} = \begin{bmatrix} 4.8 & -1.5 \\ -1.5 & 0.5 \end{bmatrix}$$

$$\boldsymbol{X}^{+} = (X^{T}X)^{-1}X^{T} = \begin{bmatrix} 4.8 & -1.5 \\ -1.5 & 0.5 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1.8 & 0.3 & -1.2 \\ -0.5 & 0 & 0.5 \end{bmatrix}$$

$$\widehat{\boldsymbol{B}} = \boldsymbol{X}^{+} \times \mathbf{Y} = \begin{bmatrix} 1.8 & 0.3 & -1.2 \\ -0.5 & 0 & 0.5 \end{bmatrix}\begin{bmatrix} 4.2 \\ 4.8 \\ 6.1 \end{bmatrix} = \begin{bmatrix} \mathbf{2.18} \\ \mathbf{0.95} \end{bmatrix} \begin{matrix} \text{切片c} \\ \text{傾きβ} \end{matrix}$$

MATLABなら簡単

Y	X		E		B
	定数	x	ε		切片c
4.2	1	2	ε1		傾きβ
4.8	1	3	ε2		
6.1	1	4	ε3		
(y=x+2, つまりβ=1, c=2を想定)					

$$\mathbf{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E}$$

$$\widehat{\boldsymbol{B}} = \boldsymbol{X}^{+} \times \mathbf{Y}$$

疑似逆行列のコマンド:**pinv**

```
>> X=[1 2; 1 3; 1 4]

X =

     1     2
     1     3
     1     4
```

```
>> Y=[4.2; 4.8; 6.1]

Y =

    4.2000
    4.8000
    6.1000
```

```
>> B=pinv(X)*Y

B =

    2.1833
    0.9500
```

$\widehat{\boldsymbol{B}}$から残差行列Eを求める

$$\mathbf{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E}$$

- Bの推定値$\widehat{\boldsymbol{B}}$

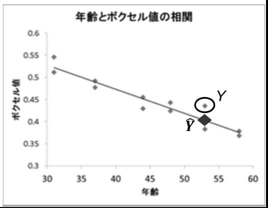
$$\widehat{\boldsymbol{B}} = \boldsymbol{X}^{+} \times \mathbf{Y}$$

- Yの予測値 $\widehat{\boldsymbol{Y}}$は直線上にある。

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X} \times \widehat{\boldsymbol{B}}$$

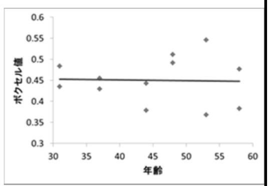
- **残差Eの推定値$\widehat{E}$は実データYと予測値$\widehat{Y}$の差分。**

$$\widehat{\boldsymbol{E}} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}}$$



t統計量の算出

- 相関解析における帰無仮説
  「年齢とボクセル値には相関がない」
  ＝「傾き**β＝0**」



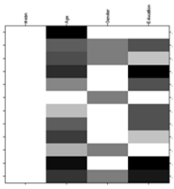
- 帰無仮説をt統計量で評価

$$\text{t統計量} = \frac{\boldsymbol{\beta}}{\boldsymbol{\beta}\text{の標準誤差}}$$

- βを算出できれば、t値が求まる
  (βの標準誤差はEから算出される)

改めてパラメータ行列Bとは

- 定数cと計画行列の各変数に呼応するβの集まり
- 定数cはBの計算では必要だが、統計量には不要（通常の統計ソフトでは、βは単独で計算される。）
- 行列Bを求めた後、各々のβを個別に取り出すための工夫が必要



$$\boldsymbol{X} \times \begin{bmatrix} c \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

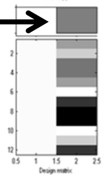
$$\boldsymbol{X} \quad \times \quad \boldsymbol{B}$$

SPMでβを取り出す方法

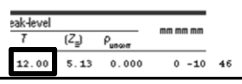
- SPMでは、パラメータ行列からβを取り出す。そのために用いるのが、コントラストの重みづけベクトル。
- パラメータ行列にコントラストベクトルをかけることで必要なパラメータだけを得ることができる。

$$\begin{pmatrix} 0 & -1 \end{pmatrix}\begin{pmatrix} c \\ \beta \end{pmatrix} = -\beta$$

コントラストベクトル　パラメータ行列



- ここで得られたβとβの標準誤差からt統計量が計算される。



群間比較（容積の比較）

- 被験者12名のうち、6名が健常（C)で6名が患者（P)
- ボクセル値をY_j、健常者の平均をβ_c、患者の平均をβ_pとすると、健常者および患者のボクセル値は次のようになる。

健常者：$Y_j = \beta_c + \varepsilon_j$

患者：$Y_j = \beta_p + \varepsilon_j$

- ダミー変数x_{c_j}およびx_{p_j}を考えると、ひとつの式であらわすことができる。

$$Y_j = \beta_c \times x_{c_j} + \beta_p \times x_{p_j} + \varepsilon_j$$

Y	Xc	Xp
0.42	1	0
0.44	1	0
0.43	1	0
0.45	1	0
0.44	1	0
0.38	1	0
0.49	0	1
0.51	0	1
0.55	0	1
0.37	0	1
0.38	0	1
0.48	0	1

ダミー変数の行列

$\boldsymbol{Y_j = \beta_c \times x_{c_j} + \beta_p \times x_{p_j} + \varepsilon_j}$ を行列で表示

$$\mathbf{Y} = \boldsymbol{X} \times \boldsymbol{B} + \boldsymbol{E}$$

Y (ボクセル値)	X (計画行列)		E (残差行列)
0.42	1	0	ε1
0.44	1	0	ε2
0.43	1	0	ε3
0.45	1	0	ε4
0.44	1	0	ε5
0.38	1	0	ε6
0.49	0	1	ε7
0.51	0	1	ε8
0.55	0	1	ε9
0.37	0	1	ε10
0.38	0	1	ε11
0.48	0	1	ε12

Y = X × B + E

B
βc
βp

パラメータ行列

β_c：健常者のボクセル値の平均
β_p：患者のボクセル値の平均

ボクセル値　計画行列　残差行列

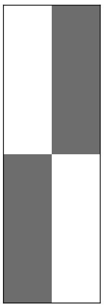
計画行列を色で表示すると

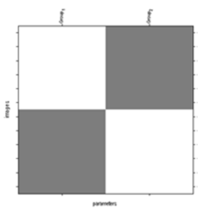
1	0
1	0
1	0
1	0
1	0
1	0
0	1
0	1
0	1
0	1
0	1
0	1

計画行列

1	0
1	0
1	0
1	0
1	0
1	0
0	1
0	1
0	1
0	1
0	1
0	1

色をつけると





SPMの画面

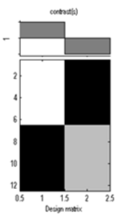
t統計量の算出

- 群間比較における帰無仮説
  「健常群と患者群の平均値には差がない」
- 帰無仮説をt統計量で評価

$$t統計量 = \frac{(Cの平均 - Pの平均)}{標準誤差}$$

$$= \frac{(\beta_c - \beta_p)}{標準誤差} = \frac{(1 \quad -1) \times \begin{pmatrix} \beta_c \\ \beta_p \end{pmatrix}}{標準誤差}$$

コントラスト



SPMの統計に対する理解を深めるために


SPM 統計入門*
Introduction to SPM statistics
Matthew Brett 著, 根本清貴訳

目次

本日の内容の大部分はこのドキュメントに載っています。SPMで画像解析をすることを考えているのならば、読んで損はしません。

www.nemotos.net/?p=41

謝辞

- 本スライドの作成にあたり、以下の先生方から多くの助言をいただきました。この場をお借りして感謝申し上げます。

- 川口淳先生（久留米大学）
- 山下典生先生（岩手医科大学）