Matrix representation

- Given the data matrix,

the data form you will see in the labs

from same unit/ object / person/ product / time / ...

Y	X_1	X_2	...	X_m
y_1	x_{11}	x_{12}	...	x_{1m}
y_2	x_{21}	x_{22}	...	x_{2m}
...		x_{ij}		
y_n	x_{n1}	x_{n2}	...	x_{nm}

ith row

x_{ij}: the observed jth variable data from ith unit

a row: one group of observations

a column: one variable (response or predictor)

- We may write a linear model as follows (*functional form*): for $i = 1, 2, \ldots, n$,

$$y_i = \beta_0 \cdot 1 + \beta_1\, g_1(x_{i1},\ldots,x_{im}) + \beta_2\, g_2(x_{i1},\ldots,x_{im}) + \ldots + \beta_{p-1}\, g_{p-1}(x_{i1},\ldots,x_{im}) + \varepsilon_i,$$

Y	**1**	g_1	g_2	...	g_{p-1}
y_1	1	g_{11}	g_{12}	...	g_{1p-1}
y_2	1	g_{21}	g_{22}	...	g_{2p-1}
...		g_{ij}			
y_n	1	g_{n1}	g_{n2}	...	g_{np-1}

a row: one group of observations

a column: response or effect

c.f.

Note. We need a model to get the matrix

where $g_{ij} = g_j(x_{i1},\ldots,x_{im})$

- ➢ the expression is (i) ugly notation (ii) conceptually awkward
- ➢ matrix/vector notation is more elegant

Y	=	β_0 × **1**	+	β_1 × g_1	+	β_2 × g_2	+	...	+	β_{p-1} × g_{p-1}	+	ε
y_1	=	1	+	g_{11}	+	g_{12}	+	...	+	g_{1p-1}	+	ε_1
y_2	=	1	+	g_{21}	+	g_{22}	+	...	+	g_{2p-1}	+	ε_2
...	=					...					+	...
y_n	=	1	+	g_{n1}	+	g_{n2}	+	...	+	g_{np-1}	+	ε_n

a row: one group of observations

a column: response or effect

c.f. → functional form

- Matrix form of the linear model:

data matrix ⟷ c.f. ⟷ model matrix

errors ε_i's are:
- uncorrelated (or independent)
- constant variance

$$\underset{n\times 1}{Y} = \underset{n\times p}{X}\,\underset{p\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

where

$$\underset{\text{random}}{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}, \quad \underset{\text{fixed}}{X} = \begin{bmatrix} 1 & g_{11} & \cdots & g_{1p-1} \\ 1 & g_{21} & \cdots & g_{2p-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & g_{n1} & \cdots & g_{np-1} \end{bmatrix}, \quad \underset{\text{fixed}}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdots \\ \beta_{p-1} \end{bmatrix}, \quad \underset{\text{random}}{\varepsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdots \\ \epsilon_n \end{bmatrix}.$$

variance-covariance matrix, cov(ε)

and $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$ and var($\boldsymbol{\varepsilon}$) $= \sigma^2 \boldsymbol{I}$, (**Note**: the assumption that errors are normally distributed is not required at the estimation stage)

$$\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \leftarrow \text{identity matrix}$$

- Example 1 (no predictor model, seen in one sample problem):

$\boxed{H_0: \mu = \mu_0 \ \text{vs.}\ H_1: \mu \neq \mu_0}$

y_i's are i.i.d. with mean μ and variance σ^2, i=1,…,n ← $\boxed{\text{statistical modeling}}$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix},\ X = \begin{bmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix},\ \boldsymbol{\beta} = \begin{bmatrix} \mu \end{bmatrix},\ \boldsymbol{\varepsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdots \\ \epsilon_n \end{bmatrix}.$$

$\boxed{\text{functional form}}$

$\varepsilon_i = y_i - \mu$

$\Rightarrow y_i = \mu + \varepsilon_i$

$E(\varepsilon_i) = 0$

$Var(\varepsilon_i) = \sigma^2$

$cov(\varepsilon_i, \varepsilon_j) = 0.$

$\boxed{\text{matrix form}}$

$Y = X\beta + \varepsilon$

$E(\varepsilon) = \underset{\sim}{0}$, $var(\varepsilon) = \sigma^2 I$

$var(Y)$ $\shortparallel$

- Example 2 (the model in two sample problem):

$\boxed{H_0: \mu_1 = \mu_2 \ \text{vs.}\ H_1: \mu_1 \neq \mu_2}$ $\boxed{\text{one predictor}}$

indep.

z_i's are i.i.d. with mean μ_1 and variance σ^2, i=1,…,m

w_j's are i.i.d. with mean μ_2 and variance σ^2, j=1,…,n

$\boxed{\text{statistical modeling}}$

$$Y = \begin{bmatrix} z_1 \\ \cdots \\ z_m \\ \hline w_1 \\ \cdots \\ w_n \end{bmatrix},\ X = \begin{array}{cc} g_1 & g_2 \\ \end{array} \begin{bmatrix} 1 & 0 \\ \cdots & \cdots \\ 1 & 0 \\ \hline 0 & 1 \\ \cdots & \cdots \\ 0 & 1 \end{bmatrix},\ \boldsymbol{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix},\ \boldsymbol{\varepsilon} = \begin{bmatrix} \epsilon_1 \\ \cdots \\ \epsilon_m \\ \delta_1 \\ \cdots \\ \delta_n \end{bmatrix}.$$

$\boxed{\text{Data matrix}}$ $\boxed{\text{model matrix}}$

Y	X	$g_1(X)$	$g_2(X)$
z_1	1	1	0
…	…	…	…
z_m	1	1	0
w_1	2	0	1
…	…	…	…
w_n	2	0	1

$\boxed{\text{functional form}}$

$\varepsilon_i = z_i - \mu_1 \Rightarrow z_i = \mu_1 + \varepsilon_i$

$\delta_j = w_j - \mu_2 \Rightarrow w_j = \mu_2 + \delta_j$

$\Rightarrow \begin{cases} E(\varepsilon) = \underset{\sim}{0} \\ var(\varepsilon) = \sigma^2 I \end{cases}$

$E(\varepsilon_i) = 0$, $E(\delta_j) = 0$

$var(\varepsilon_i) = var(\delta_j) = \sigma^2$

$cov(\varepsilon_i, \varepsilon_{i'})$

$= cov(\delta_j, \delta_{j'})$

$= cov(\varepsilon_i, \delta_j) = 0$

$\boxed{\text{matrix form}}$

$Y = \mu_1 g_1 + \mu_2 g_2 + \varepsilon$

$= X\beta + \varepsilon$

$E(\varepsilon) = \underset{\sim}{0}$, $var(\varepsilon) = \sigma^2 I$

$\| = var(Y)$

❖ **Reading**: Faraway (2005, 1st ed.), 2.2

❖ **Further reading**: D&S, 4.1