SPM and some of its Maths
An Introduction

Joana Leitão

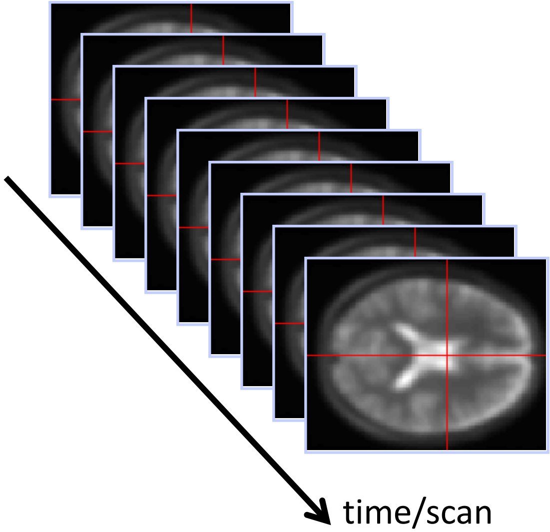
MRes Advanced Brain Imaging, 2013

Linear Algebra – What is a matrix?

- System of equations are very common in science and mathematics
- Matrices are a compact and convenient way of writing down systems of linear equations
- They allow for computational implementations

Time-series in one single voxel

$$\begin{aligned}
Y_1 &= x_{11} * \beta_1 + \cdots + x_{1i} * \beta_i + \cdots + x_{1I} * \beta_I + \varepsilon_1 \\
\vdots \quad & \quad \vdots \qquad\qquad\quad \vdots \qquad\qquad\quad \vdots \qquad\quad \vdots \\
Y_j &= x_{j1} * \beta_1 + \cdots + x_{ji} * \beta_i + \cdots + x_{jI} * \beta_I + \varepsilon_j \\
\vdots \quad & \quad \vdots \qquad\qquad\quad \vdots \qquad\qquad\quad \vdots \qquad\quad \vdots \\
Y_J &= x_{J1} * \beta_1 + \cdots + x_{Ji} * \beta_i + \cdots + x_{JI} * \beta_I + \varepsilon_J
\end{aligned}$$



Y_j – *intensity in one voxel at time/scan j*

x_{ji} – *value of regressor i at time/scan j*

β_i – *value of beta parameter for regressor i*

ε_j – *error term associated with the jth scan*

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{ccccccccccc}
Y_1 & = & x_{11} * \beta_1 & + \cdots + & x_{1i} * \beta_i & + \cdots + & x_{1I} * \beta_I & + & \varepsilon_1 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
Y_j & = & x_{j1} * \beta_1 & + \cdots + & x_{ji} * \beta_i & + \cdots + & x_{jI} * \beta_I & + & \varepsilon_j \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
Y_J & = & x_{J1} * \beta_1 & + \cdots + & x_{Ji} * \beta_i & + \cdots + & x_{JI} * \beta_I & + & \varepsilon_J
\end{array}$$

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$
\begin{array}{|c|lllll}
\cline{1-1}
Y_1 & = x_{11} * \beta_1 + \cdots + & x_{1i} * \beta_i + \cdots + & x_{1I} * \beta_I + & \varepsilon_1 \\
\vdots & \quad \vdots & \quad \vdots & \quad \vdots & \vdots \\
Y_j & = x_{j1} * \beta_1 + \cdots + & x_{ji} * \beta_i + \cdots + & x_{jI} * \beta_I + & \varepsilon_j \\
\vdots & \quad \vdots & \quad \vdots & \quad \vdots & \vdots \\
Y_J & = x_{J1} * \beta_1 + \cdots + & x_{Ji} * \beta_i + \cdots + & x_{JI} * \beta_I + & \varepsilon_J \\
\cline{1-1}
\end{array}
$$

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\boxed{\begin{matrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{matrix}} \begin{matrix} = x_{11} * \beta_1 + \cdots + x_{1i} * \beta_i + \cdots + x_{1I} * \beta_I + \varepsilon_1 \\ \vdots \qquad\qquad\quad \vdots \qquad\qquad\qquad \vdots \qquad\qquad\quad \vdots \\ = x_{j1} * \beta_1 + \cdots + x_{ji} * \beta_i + \cdots + x_{jI} * \beta_I + \varepsilon_j \\ \vdots \qquad\qquad\quad \vdots \qquad\qquad\qquad \vdots \qquad\qquad\quad \vdots \\ = x_{J1} * \beta_1 + \cdots + x_{Ji} * \beta_i + \cdots + x_{JI} * \beta_I + \varepsilon_J \end{matrix}$$

$$\boxed{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} =$$

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\boxed{\begin{array}{c} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{array}} \begin{array}{lllll} = x_{11} * \beta_1 + \cdots + & x_{1i} * \beta_i + \cdots + & x_{1I} * \beta_I + & \varepsilon_1 \\ \quad \vdots & \quad \vdots & \quad \vdots & \vdots \\ = x_{j1} * \beta_1 + \cdots + & x_{ji} * \beta_i + \cdots + & x_{jI} * \beta_I + & \varepsilon_j \\ \quad \vdots & \quad \vdots & \quad \vdots & \vdots \\ = x_{J1} * \beta_1 + \cdots + & x_{Ji} * \beta_i + \cdots + & x_{JI} * \beta_I + & \varepsilon_J \end{array}$$

$$\boxed{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} =$$

vector
(Jx1)

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{ccccccccc}
Y_1 & = x_{11} * \beta_1 & + \cdots + & x_{1i} * \beta_i & + \cdots + & x_{1I} * \beta_I & + & \varepsilon_1 \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
Y_j & = x_{j1} * \beta_1 & + \cdots + & x_{ji} * \beta_i & + \cdots + & x_{jI} * \beta_I & + & \varepsilon_j \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
Y_J & = x_{J1} * \beta_1 & + \cdots + & x_{Ji} * \beta_i & + \cdots + & x_{JI} * \beta_I & + & \varepsilon_J
\end{array}$$

$$\boxed{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} =$$

vector

(Jx1)

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{c}Y_1 = x_{11} * \beta_1 + \cdots + x_{1i} * \beta_i + \cdots + x_{1I} * \beta_I + \varepsilon_1 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ Y_j = x_{j1} * \beta_1 + \cdots + x_{ji} * \beta_i + \cdots + x_{jI} * \beta_I + \varepsilon_j \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ Y_J = x_{J1} * \beta_1 + \cdots + x_{Ji} * \beta_i + \cdots + x_{JI} * \beta_I + \varepsilon_J\end{array}$$

$$\boxed{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} = \qquad\qquad\qquad + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}}$$

vector
(Jx1)

vector
(Jx1)

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$
\begin{array}{c}
Y_1 = x_{11} * \boxed{\beta_1} + \cdots + x_{1i} * \boxed{\beta_i} + \cdots + x_{1I} * \boxed{\beta_I} + \varepsilon_1 \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
Y_j = x_{j1} * \boxed{\beta_1} + \cdots + x_{ji} * \boxed{\beta_i} + \cdots + x_{jI} * \boxed{\beta_I} + \varepsilon_j \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
Y_J = x_{J1} * \boxed{\beta_1} + \cdots + x_{Ji} * \boxed{\beta_i} + \cdots + x_{JI} * \boxed{\beta_I} + \varepsilon_J
\end{array}
$$

$$
\boxed{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} = \qquad\qquad\qquad\qquad + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}}
$$

vector
(Jx1)

vector
(Jx1)

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{ccccccccccc}
Y_1 & = x_{11} * & \beta_1 & + \cdots + x_{1i} * & \beta_i & + \cdots + x_{1I} * & \beta_I & + & \varepsilon_1 \\
\vdots & \vdots & & \vdots & & \vdots & & & \vdots \\
Y_j & = x_{j1} * & \beta_1 & + \cdots + x_{ji} * & \beta_i & + \cdots + x_{jI} * & \beta_I & + & \varepsilon_j \\
\vdots & \vdots & & \vdots & & \vdots & & & \vdots \\
Y_J & = x_{J1} * & \beta_1 & + \cdots + x_{Ji} * & \beta_i & + \cdots + x_{JI} * & \beta_I & + & \varepsilon_J
\end{array}$$

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix} = \qquad\qquad \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_I \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}$$

vector
(Jx1)

vector
(Ix1)

vector
(Jx1)

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{l}
Y_1 = x_{11} * \beta_1 + \cdots + x_{1i} * \beta_i + \cdots + x_{1I} * \beta_I + \varepsilon_1 \\
\vdots \\
Y_j = x_{j1} * \beta_1 + \cdots + x_{ji} * \beta_i + \cdots + x_{jI} * \beta_I + \varepsilon_j \\
\vdots \\
Y_J = x_{J1} * \beta_1 + \cdots + x_{Ji} * \beta_i + \cdots + x_{JI} * \beta_I + \varepsilon_J
\end{array}$$

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix} = \qquad \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_I \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}$$

vector
(Jx1)

vector
(Ix1)

vector
(Jx1)

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$
\begin{array}{c}
Y_1 = \boxed{x_{11}} * \boxed{\beta_1} + \cdots + \boxed{x_{1i}} * \boxed{\beta_i} + \cdots + \boxed{x_{1I}} * \boxed{\beta_I} + \varepsilon_1 \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
Y_j = \boxed{x_{j1}} * \boxed{\beta_1} + \cdots + \boxed{x_{ji}} * \boxed{\beta_i} + \cdots + \boxed{x_{jI}} * \boxed{\beta_I} + \varepsilon_j \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
Y_J = \boxed{x_{J1}} * \boxed{\beta_1} + \cdots + \boxed{x_{Ji}} * \boxed{\beta_i} + \cdots + \boxed{x_{JI}} * \boxed{\beta_I} + \varepsilon_J
\end{array}
$$

$$
\boxed{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} = \boxed{\begin{bmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1I} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{ji} & \cdots & x_{jI} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{J1} & \cdots & x_{Ji} & \cdots & x_{JI} \end{bmatrix}} \boxed{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_I \end{bmatrix}} + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}}
$$

vector (Jx1) matrix (JxI) vector (Ix1) vector (Jx1)

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{c}
Y_1 = x_{11} * \beta_1 + \cdots + x_{1i} * \beta_i + \cdots + x_{1I} * \beta_I + \varepsilon_1 \\
\vdots \\
Y_j = x_{j1} * \beta_1 + \cdots + x_{ji} * \beta_i + \cdots + x_{jI} * \beta_I + \varepsilon_j \\
\vdots \\
Y_J = x_{J1} * \beta_1 + \cdots + x_{Ji} * \beta_i + \cdots + x_{JI} * \beta_I + \varepsilon_J
\end{array}$$

$$\underbrace{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}}_{\substack{\text{vector} \\ \text{(Jx1)}}} = \underbrace{\begin{bmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1I} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{ji} & \cdots & x_{jI} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{J1} & \cdots & x_{Ji} & \cdots & x_{JI} \end{bmatrix}}_{\substack{\text{matrix} \\ \text{(JxI)}}} * \underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_I \end{bmatrix}}_{\substack{\text{vector} \\ \text{(Ix1)}}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}}_{\substack{\text{vector} \\ \text{(Jx1)}}}$$

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{c}
Y_1 = x_{11} * \beta_1 + \cdots + x_{1i} * \beta_i + \cdots + x_{1I} * \beta_I + \varepsilon_1 \\
\vdots \\
Y_j = x_{j1} * \beta_1 + \cdots + x_{ji} * \beta_i + \cdots + x_{jI} * \beta_I + \varepsilon_j \\
\vdots \\
Y_J = x_{J1} * \beta_1 + \cdots + x_{Ji} * \beta_i + \cdots + x_{JI} * \beta_I + \varepsilon_J
\end{array}$$

$$\underset{\substack{\text{vector} \\ \text{(Jx1)}}}{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} = \underset{\substack{\text{matrix} \\ \text{(JxI)}}}{\begin{bmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1I} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{ji} & \cdots & x_{jI} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{J1} & \cdots & x_{Ji} & \cdots & x_{JI} \end{bmatrix}} * \underset{\substack{\text{vector} \\ \text{(Ix1)}}}{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_I \end{bmatrix}} + \underset{\substack{\text{vector} \\ \text{(Jx1)}}}{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}}$$

$$\boxed{y = X\beta + e}$$

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{array}{ccccccccccccccc}
Y_1 & = & \boxed{x_{11}} & * & \boxed{\beta_1} & + \cdots + & \boxed{x_{1i}} & * & \boxed{\beta_i} & + \cdots + & \boxed{x_{1I}} & * & \boxed{\beta_I} & + & \varepsilon_1 \\
\vdots & & \vdots & & & & \vdots & & & & \vdots & & & & \vdots \\
Y_j & = & \boxed{x_{j1}} & * & \boxed{\beta_1} & + \cdots + & \boxed{x_{ji}} & * & \boxed{\beta_i} & + \cdots + & \boxed{x_{jI}} & * & \boxed{\beta_I} & + & \varepsilon_j \\
\vdots & & \vdots & & & & \vdots & & & & \vdots & & & & \vdots \\
Y_J & = & \boxed{x_{J1}} & * & \boxed{\beta_1} & + \cdots + & \boxed{x_{Ji}} & * & \boxed{\beta_i} & + \cdots + & \boxed{x_{JI}} & * & \boxed{\beta_I} & + & \varepsilon_J
\end{array}$$

- Matrices are usually represented by capital bold letters
- Vectors are usually represented by small bold letters

$$\begin{array}{ccccccc}
\boxed{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} & = & \boxed{\begin{bmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1I} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{ji} & \cdots & x_{jI} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{J1} & \cdots & x_{Ji} & \cdots & x_{JI} \end{bmatrix}} & * & \boxed{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_I \end{bmatrix}} & + & \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}} \\
\text{vector} & & \text{matrix} & & \text{vector} & & \text{vector} \\
\text{(Jx1)} & & \text{(JxI)} & & \text{(Ix1)} & & \text{(Jx1)}
\end{array}$$

$$\boxed{y = X\beta + e}$$

Note: if J = I the matrix is said to be squared

Linear Algebra – What is a matrix?

Time-series in one single voxel

$$\begin{aligned} Y_1 &= x_{11} * \beta_1 + \cdots + x_{1i} * \beta_i + \cdots + x_{1I} * \beta_I + \varepsilon_1 \\ &\vdots \\ Y_j &= x_{j1} * \beta_1 + \cdots + x_{ji} * \beta_i + \cdots + x_{jI} * \beta_I + \varepsilon_j \\ &\vdots \\ Y_J &= x_{J1} * \beta_1 + \cdots + x_{Ji} * \beta_i + \cdots + x_{JI} * \beta_I + \varepsilon_J \end{aligned}$$

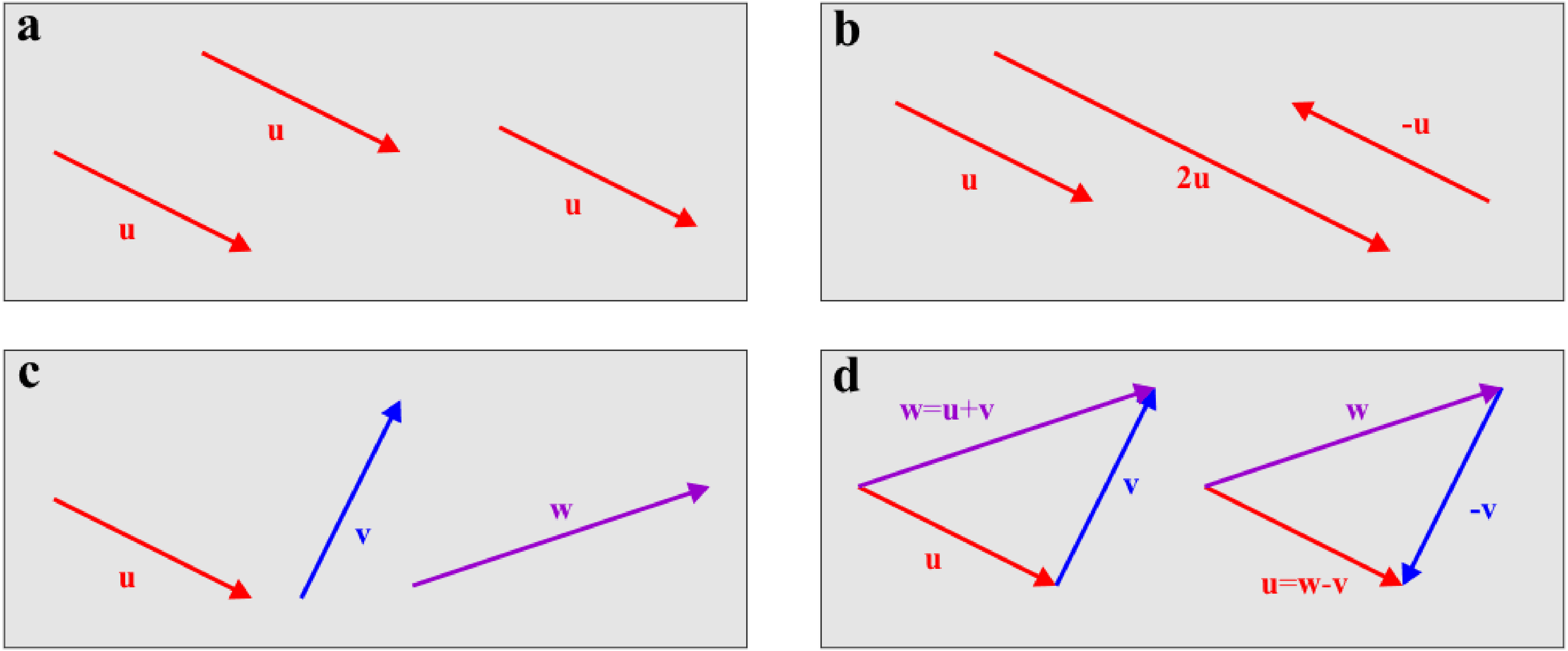
- Matrices are usually represented by capital bold letters
- Vectors are usually represented by small bold letters

$$\underset{\substack{\text{vector}\\ \text{(Jx1)}}}{\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{bmatrix}} = \underset{\substack{\text{matrix}\\ \text{(JxI)}}}{\begin{bmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1I} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{ji} & \cdots & x_{jI} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{J1} & \cdots & x_{Ji} & \cdots & x_{JI} \end{bmatrix}} * \underset{\substack{\text{vector}\\ \text{(Ix1)}}}{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_I \end{bmatrix}} + \underset{\substack{\text{vector}\\ \text{(Jx1)}}}{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_j \\ \vdots \\ \varepsilon_J \end{bmatrix}}$$

$$\boxed{y = X\beta + e}$$

Note: if J = I the matrix is said to be squared

A note on vectors – geometric perspective



Vectors. **a**: identical vectors at different locations. **b**: scaled vectors (linearly dependent). **c**: different vectors (linearly independent). **d**: vector addition and subtraction

Linear Algebra – Matrix Addition

Note:

1) matrices have to have the same dimensions

2) you just have to add the corresponding entries

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}+\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$=\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

Linear Algebra – Matrix Multiplication

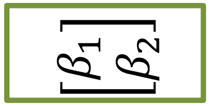
A smaller example

$$\begin{aligned} Y_1 &= x_{11} * \beta_1 + x_{12} * \beta_2 \\ Y_2 &= x_{21} * \beta_1 + x_{22} * \beta_2 \\ Y_3 &= x_{31} * \beta_1 + x_{32} * \beta_2 \end{aligned} \quad \Longrightarrow \quad \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

Multiplication from (nxm)-matrix with (mx1)-vector = (nx1)-vector

For the ith entry of the resulting vector:

1) tilt the **green** vector



$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} * \boxed{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}$$

2) multiply overlapping entries of the ith row of the matrix and vector

   - e.g. for i = 1: $x_{11} * \beta_1; \; x_{12} * \beta_2$

3) Add them: $Y_{11} = x_{11} * \beta_1 + x_{12} * \beta_2$

Note: number of rows in the vector has to be the same as the number of columns in the matrix

Linear Algebra – Matrix Multiplication

Multiplication from (mxn)-matrix with (nxk)-matrix = (mxk)-matrix

For the (i,j) entry of the resulting matrix:

1) tilt the jth column of the second matrix

   - e.g. for entry (2,1):

$$\boxed{\begin{bmatrix}\beta_{11} & \beta_{21}\end{bmatrix}}$$

$$\begin{bmatrix}x_{11} & x_{12}\\ x_{21} & x_{22}\\ x_{31} & x_{32}\end{bmatrix} * \begin{bmatrix}\boxed{\begin{matrix}\beta_{11}\\ \beta_{21}\end{matrix}} & \begin{matrix}\beta_{12}\\ \beta_{22}\end{matrix}\end{bmatrix}$$

2) multiply overlapping entries of the ith row of the first matrix and the tilted vector

   - e.g. for entry (2,1): $x_{21} * \beta_{11};\ x_{22} * \beta_{21}$

3) add them:

   - e.g. for entry (2,1): $Y_{21} = x_{21} * \beta_{11} + x_{22} * \beta_{21}$

Note: number of rows in the second matrix has to be the same as the number of columns in the matrix

Linear Algebra – Matrix Multiplication

Properties of Matrix Multiplication

1) **Not** commutative (**!**):

$$\underbrace{\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} * \begin{bmatrix} \beta_{11}\beta_{12}\beta_{13} \\ \beta_{21}\beta_{22}\beta_{23} \end{bmatrix}}_{\text{(3x3) - matrix}} \neq \underbrace{\begin{bmatrix} \beta_{11}\beta_{12}\beta_{13} \\ \beta_{21}\beta_{22}\beta_{23} \end{bmatrix} * \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}}_{\text{(4x4) - matrix}}$$

Linear Algebra – Matrix Multiplication

Properties of Matrix Multiplication

1) Associative: $\boldsymbol{A} * (\boldsymbol{B} * \boldsymbol{C}) = (\boldsymbol{A} * \boldsymbol{B}) * \boldsymbol{C}$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} * \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \right)$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} (b_{11}c_{11} + b_{12}c_{21}) & (b_{11}c_{12} + b_{12}c_{22}) \\ (b_{21}c_{11} + b_{22}c_{21}) & (b_{21}c_{12} + b_{22}c_{22}) \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11}c_{11} + a_{11}b_{12}c_{21} + a_{12}b_{21}c_{11} + a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12} + a_{11}b_{12}c_{22} + a_{12}b_{21}c_{12} + a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11} + a_{21}b_{12}c_{21} + a_{22}b_{21}c_{11} + a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12} + a_{21}b_{12}c_{22} + a_{22}b_{21}c_{12} + a_{22}b_{22}c_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} * (b_{11}c_{11} + b_{12}c_{21}) + a_{12} * (b_{21}c_{11} + b_{22}c_{21}) & a_{11} * (b_{11}c_{12} + b_{12}c_{22}) + a_{12} * (b_{21}c_{12} + b_{22}c_{22}) \\ a_{21} * (b_{11}c_{11} + b_{12}c_{21}) + a_{22} * (b_{21}c_{11} + b_{22}c_{21}) & a_{21} * (b_{11}c_{12} + b_{12}c_{22}) + a_{22} * (b_{21}c_{12} + b_{22}c_{22}) \end{bmatrix}$$

$$= \begin{bmatrix} (a_{11}b_{11}+a_{12}b_{21}) * c_{11} + (a_{11}b_{12}+a_{12}b_{22}) * c_{21} & (a_{21}b_{11}+a_{22}b_{21}) * c_{12} + (a_{21}b_{12}+a_{22}b_{22}) * c_{22} \\ (a_{21}b_{11}+a_{22}b_{21}) * c_{11} + (a_{21}b_{12}+a_{22}b_{22}) * c_{21} & (a_{21}b_{11}+a_{22}b_{21}) * c_{12} + (a_{21}b_{12}+a_{22}b_{22}) * c_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} * \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

$$= \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right) * \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = (\boldsymbol{A} * \boldsymbol{B}) * \boldsymbol{C}$$

Linear Algebra – Matrix Multiplication

Properties of Matrix Multiplication

2) Distributive: $\boldsymbol{A} * (\boldsymbol{B} + \boldsymbol{C}) \; = \; (\boldsymbol{A} * \boldsymbol{B}) + (\boldsymbol{A} * \boldsymbol{C})$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \right) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} b_{11} + c_{11} & b_{12} + c_{12} \\ b_{21} + c_{21} & b_{22} + c_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} * (b_{11} + c_{11}) + a_{12} * (b_{21} + c_{21}) & a_{11} * (b_{12} + c_{12}) + a_{12} * (b_{22} + c_{22}) \\ a_{21} * (b_{11} + c_{11}) + a_{22} * (b_{21} + c_{21}) & a_{21} * (b_{12} + c_{12}) + a_{22} * (b_{22} + c_{22}) \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{11}c_{11} + a_{12}c_{21} & a_{11}b_{12} + a_{12}b_{22} + a_{11}c_{12} + a_{12}c_{22} \\ a_{21}b_{11} + a_{22}b_{21} + a_{21}c_{11} + a_{22}c_{21} & a_{21}b_{12} + a_{22}b_{22} + a_{21}c_{12} + a_{22}c_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} * \begin{bmatrix} a_{11}c_{11} + a_{12}c_{21} & a_{11}c_{12} + a_{12}c_{22} \\ a_{21}c_{11} + a_{22}c_{21} & a_{21}c_{12} + a_{22}c_{22} \end{bmatrix}$$

$$= \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right) + \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \right) = (\boldsymbol{A} * \boldsymbol{B}) + (\boldsymbol{A} * \boldsymbol{C})$$

Linear Algebra – Transpose

For some matrix $\boldsymbol{X}$ defined by:

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}$$

its transpose $\boldsymbol{X}^T$ is given by:

$$\boldsymbol{X}^T = \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \end{bmatrix}$$

Basically you just switch rows with columns.

In MATLAB:

If X is a matrix its transpose is given by X’

Linear Algebra – Transpose

Some properties:

1) $(\boldsymbol{A}^T)^T = \boldsymbol{A}$

2) $(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$

$$\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\right)^T = \left(\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}\right)^T$$

$$= \left(\begin{bmatrix} a_{11}+b_{11} & a_{21}+b_{21} \\ a_{12}+b_{12} & a_{22}+b_{22} \end{bmatrix}\right)$$

$$= \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} = \boldsymbol{A}^T + \boldsymbol{B}^T$$

Linear Algebra – Transpose

Some properties:

3) $(\boldsymbol{A} * \boldsymbol{B})^{\boldsymbol{T}} = \boldsymbol{B}^{\boldsymbol{T}} * \boldsymbol{A}^{\boldsymbol{T}}$

$$(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix})^T = (\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix})^T$$

$$= \left(\begin{bmatrix} b_{11}a_{11} + b_{21}a_{12} & b_{11}a_{21} + b_{21}a_{22} \\ b_{12}a_{11} + b_{22}a_{12} & b_{12}a_{21} + b_{22}a_{22} \end{bmatrix}\right)$$

$$= \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} * \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = \boldsymbol{B}^{\boldsymbol{T}} * \boldsymbol{A}^{\boldsymbol{T}}$$

Note: If $\boldsymbol{a}$ and $\boldsymbol{b}$ are two vectors of size $(nx1)$, then:

$$\boldsymbol{a}^T * \boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i = \boldsymbol{a} \cdot \boldsymbol{b}$$ **(inner product)**

Hence for a vector $\boldsymbol{v}$ of entries v_i , with $i = 1, \ldots N$ we have $\sum_{i=1}^{N} {v_i}^2 = \boldsymbol{v}^T * \boldsymbol{v}$

Linear Algebra – Identity and Zero Matrices

Identity matrix:

$$I_1 = \begin{bmatrix}1\end{bmatrix},\ I_2 = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\ I_3 = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix},\ \cdots,\ I_n = \begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

Note: 1) The identity matrix is always squared

2) The identity matrix is **diagonal**, i.e. its only non-zero entries are in the main diagonal

Zero matrix is a (square) matrix which entries all equal zero.

In MATLAB:

Identity matrix of dimensions mxm: eye(m)

Identity matrix of dimensions mxm: zeros(m, m)

Linear (In)dependence

A collection of vectors is a **linearly independent** if none of them can be written as a linear combination of finitely many other vectors in the collection.

A collection of vectors which is not linearly independent is called **linearly dependent**.

Linear (In)dependence

A collection of vectors is a **linearly independent** if none of them can be written as a linear combination of finitely many other vectors in the collection.

A collection of vectors which is not linearly independent is called **linearly dependent**.

Linearly Dependence:

The $(Ix1)$-vector $\boldsymbol{y}$ is said to be a linearly dependent to the $(Ix1)$-vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$ if there exist non-zero constants c_i such that

$$\boldsymbol{y} = \sum_{i=1}^{N} c_i \boldsymbol{x}_i$$

Linear (In)dependence

A collection of vectors is a **linearly independent** if none of them can be written as a linear combination of finitely many other vectors in the collection.

A collection of vectors which is not linearly independent is called **linearly dependent**.

Example:

$$\underbrace{\overbrace{\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}0\\2\\-2\end{bmatrix},\begin{bmatrix}1\\-2\\1\end{bmatrix}}^{\text{independent}},\begin{bmatrix}4\\2\\3\end{bmatrix}}_{\text{dependent}}$$

Linearly Dependence:

The $(Ix1)$-vector $\boldsymbol{y}$ is said to be a linearly dependent to the $(Ix1)$-vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$ if there exist non-zero constants c_i such that

$$\boldsymbol{y} = \sum_{i=1}^{N} c_i \boldsymbol{x}_i$$

Linear (In)dependence

A collection of vectors is a **linearly independent** if none of them can be written as a linear combination of finitely many other vectors in the collection.

A collection of vectors which is not linearly independent is called **linearly dependent**.

Example:

$$\underbrace{\overbrace{\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}0\\2\\-2\end{bmatrix},\begin{bmatrix}1\\-2\\1\end{bmatrix}}^{\text{independent}},\begin{bmatrix}4\\2\\3\end{bmatrix}}_{\text{dependent}}$$

Linearly Dependence:

The $(Ix1)$-vector $\boldsymbol{y}$ is said to be a linearly dependent to the $(Ix1)$-vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$ if there exist non-zero constants c_i such that

$$\boldsymbol{y} = \sum_{i=1}^{N} c_i \boldsymbol{x}_i$$

A square matrix is called **non-singular** if its columns form a linear independent set. Otherwise it is called **singular**.

Linear Algebra – Inverse Matrix

A (nxn)-matrix $\boldsymbol{A}$ is called invertible if there exists an (nxn)- matrix $\boldsymbol{B}$ such that:

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I_n}$$

There are several methods to obtain the inverse of a squared matrix, as for example the Gauss-Jordan method.
Note that not all matrices are invertible. If a matrix has linearly dependent rows or columns it is not invertible.

In MATLAB:

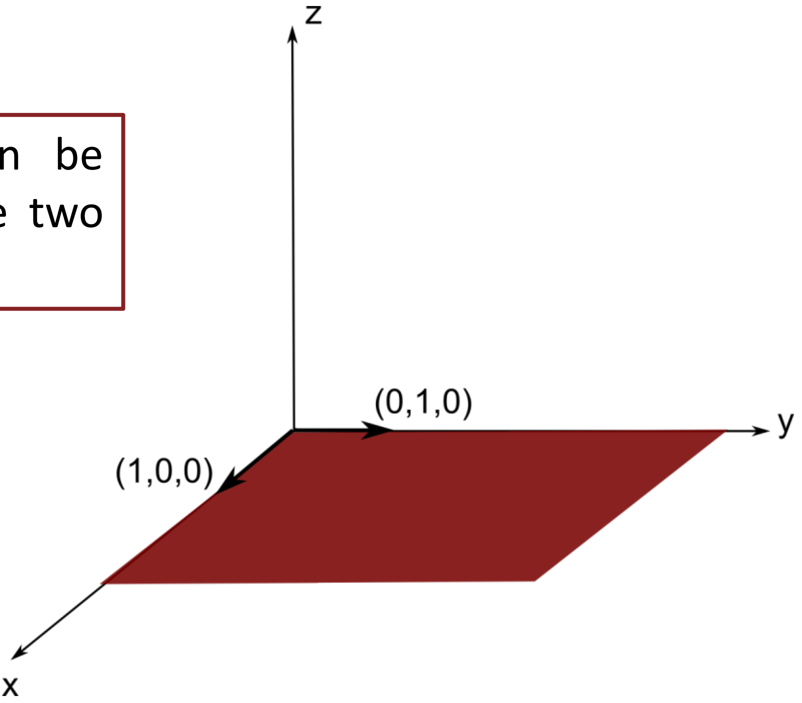
If $\boldsymbol{X}$ is a matrix its inverse is given by inv($\boldsymbol{X}$)

Linear Algebra – Span of a space

If you have a collection of linearly independent vectors, geometrically this means that they are not co-linear. Hence they span a certain space which consists of **all possible linear combinations** that you can have with these vectors.

Consider for example the two linear independent vectors in 3D space (1,0,0) and (0,1,0). These vectors span a 2D space (actually the xy-plane) in a 3D space.

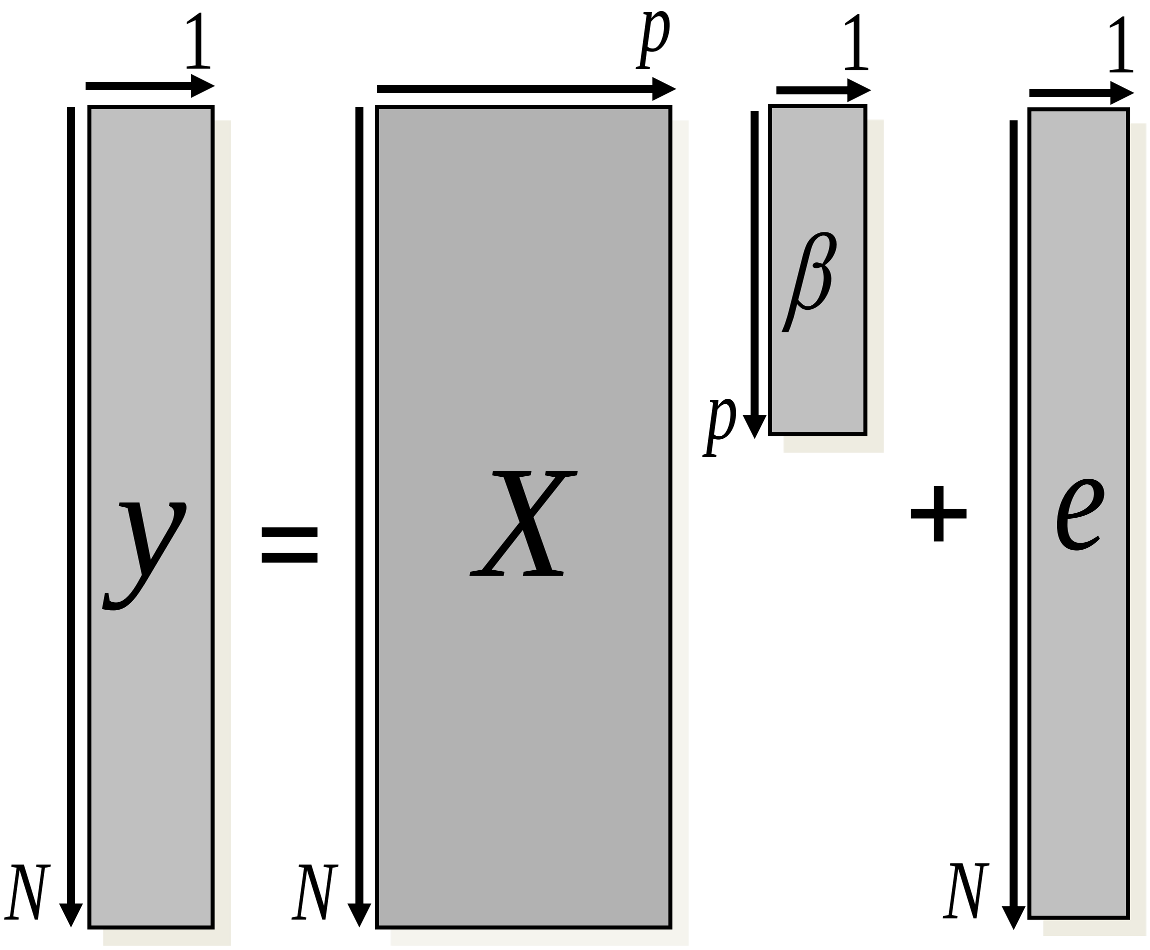
Indeed, all points in the xy-plane can be written as a linear combination of these two vectors



CHANGE OF TOPIC!

Good time for a break! ☺

GLM



$$y = X\beta + e$$

$$e \sim N(0, \sigma^2 I)$$

Model is specified by

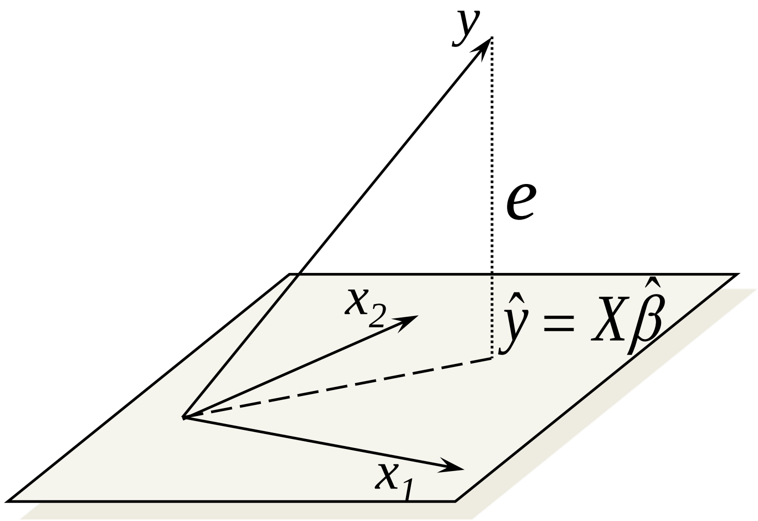
1. Design matrix X
2. Assumptions about e

N: number of scans

p: number of regressors

The design matrix embodies all available knowledge about experimentally controlled factors and potential confounds.

Projection matrix & Ordinary Least Squares

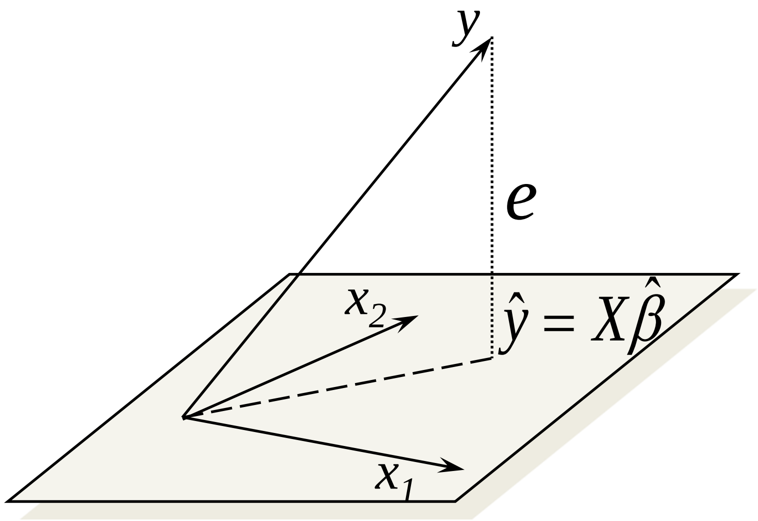


Back to our initial example:

$$\boxed{y = X\beta + e}$$

If $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are the columns (i.e., your regressors) of your design matrix $\boldsymbol{X}$ they will span a certain space. Generally, the data you acquire $\boldsymbol{y}$ will not be in the space spanned by your regressors. In other words, your data cannot be completely explained by your design matrix since there is always some error involved. The aim is to find the point $\widehat{\boldsymbol{y}}$ in the space spanned by your regressors that minimizes this error.

Projection matrix & Ordinary Least Squares



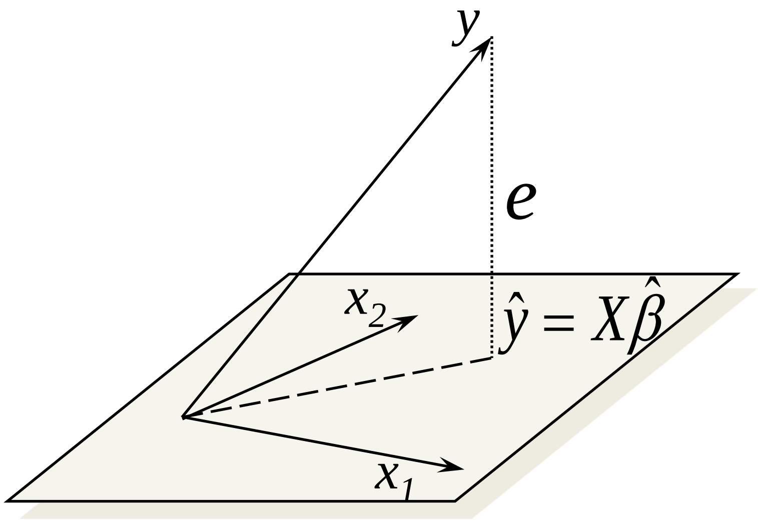
Back to our initial example:

$$\boxed{y = X\beta + e}$$

If $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are the columns (i.e., your regressors) of your design matrix $\boldsymbol{X}$ they will span a certain space. Generally, the data you acquire $\boldsymbol{y}$ will not be in the space spanned by your regressors. In other words, your data cannot be completely explained by your design matrix since there is always some error involved. The aim is to find the point $\widehat{\boldsymbol{y}}$ in the space spanned by your regressors that minimizes this error.

Since $\widehat{\boldsymbol{y}}$ is in the space spanned by you regressors, it can be written as a linear combination of those. Hence, finding this point is equivalent to finding by how much you have to multiply $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively. These constants/values are what the $\boldsymbol{\beta}$s represent.

Projection matrix & Ordinary Least Squares



Back to our initial example:

$$\boxed{y = X\beta + e}$$

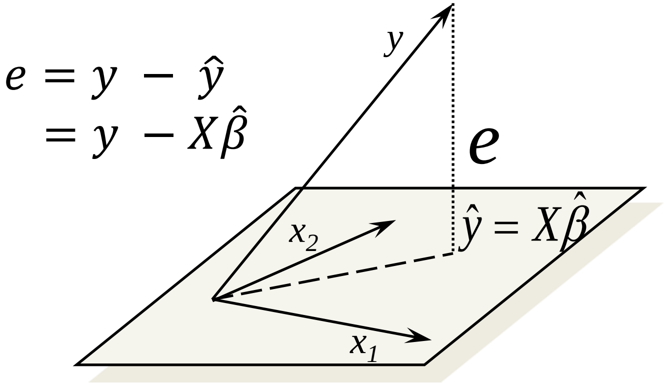
Note: If you have a plane and a point, the shortest distance between them is defined by the line passing through the point and perpendicular to the plane. In other words, the minimal error $\boldsymbol{e}$ is given by the projection of $\boldsymbol{y}$ onto the space spanned by your regressors.

The **Least squares** is a method that gives an approximate solution for these *overdetermined* (in other words, the number of explanatory variables you have is less than the observed data) systems. Instead of minimizing the sum of the errors, the method of least squares minimizes sum of the squared errors.

Ordinary Least Squares Derived

So the aim is to minimize:

$$\sum_{j=1}^{J} {e_j}^2 = \boldsymbol{e}^{\boldsymbol{T}} * \boldsymbol{e}$$



This is the inner product of $\boldsymbol{e}$ by itself and is hence a scalar:

$$\sum_{j=1}^{J} {e_j}^2 = \boldsymbol{e}^{\boldsymbol{T}} * \boldsymbol{e}$$

$$= \left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)^{\boldsymbol{T}} * \left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) = \left(\boldsymbol{y}^{\boldsymbol{T}} - \widehat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\right) * \left(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)$$

$$= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - \underbrace{\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{X}\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\mathbf{y}} + \widehat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

Both multiplications involve the same vectors/matrices, so the same numbers/entries are involved. Since they are scalar, they are actually the same so we can just add them together

Just try with a small example to see that's true.

$$= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - 2\widehat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\mathbf{y} + \widehat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

In the last equation we know everything except for the $\boldsymbol{\beta}$, which is what we want to find out. Hence $\boldsymbol{e}$ is a scalar that depends on all the betas . In other words it is a function of the betas.

Ordinary Least Squares Derived

Since $\sum_{j=1}^{J} e_j{}^2$ is a function of the betas, to minimize it implies to differentiate with respect to every single β (i.e., to the vector of the betas) and set this derivative to zero.

Differentiate with respect to a vector means you differentiate with respect to each component of the vector:

$$\frac{\partial S}{\partial \beta} = \begin{bmatrix} \frac{\partial S}{\partial \beta_1} \\ \vdots \\ \frac{\partial S}{\partial \beta_I} \end{bmatrix}$$

Ordinary Least Squares Derived

$$\boxed{\frac{\partial S}{\partial \beta} = \begin{bmatrix} \frac{\partial S}{\partial \beta_1} \\ \vdots \\ \frac{\partial S}{\partial \beta_I} \end{bmatrix}}$$

Let’s differentiate it by parts:

$$S = \sum_{j=1}^{J} {e_j}^2 = \underbrace{\boldsymbol{y}^T\boldsymbol{y}}_{\text{Part I}} - \underbrace{\mathbf{2}\widehat{\boldsymbol{\beta}}^T\boldsymbol{X}^T\mathbf{y}}_{\text{Part II}} + \underbrace{\widehat{\boldsymbol{\beta}}^T\boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}}}_{\text{Part III}}$$

1) Part I: $\boldsymbol{y}^T\boldsymbol{y}$ does not depend on any β so its derivative is zero.

2) Part II: To simplify, let’s consider a small example. Suppose you have two regressors and only 3 data points. Hence, $\boldsymbol{X}$ is a $(3x2)$-matrix, $\boldsymbol{y}$ is a $(3x1)$-vector and $\boldsymbol{\beta}$ is a $(2x1)$-vector. Let’s also say that $\mathbf{A} = \boldsymbol{X}^T\mathbf{y}$ $((2x1)$-vector$)$.

$$-2\widehat{\boldsymbol{\beta}}^T\boldsymbol{X}^T\mathbf{y} = -2\widehat{\boldsymbol{\beta}}^T\boldsymbol{A} = -2 * \begin{bmatrix}\beta_1 & \beta_2\end{bmatrix} * \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$$
$$= -2 * (a_{11}\beta_1 + a_{21}\beta_2)$$

Derivative with respect to β_1: $-2a_{11}$
Derivative with respect to β_2: $-2a_{21}$

⇒ Derivative $= -2 * \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$

$$= -2\mathbf{A} = -2\boldsymbol{X}^T\mathbf{y}$$

Ordinary Least Squares Derived

Let's differentiate it by parts:

$$S = \sum_{j=1}^{J} {e_j}^2 = \underbrace{\boldsymbol{y}^T\boldsymbol{y}}_{\text{Part I}} - \underbrace{\boldsymbol{2\hat{\beta}^T X^T}\mathbf{y}}_{\text{Part II}} + \underbrace{\boldsymbol{\hat{\beta}^T X^T X\hat{\beta}}}_{\text{Part III}}$$

$$\boxed{\frac{\partial S}{\partial \beta} = \begin{bmatrix} \frac{\partial S}{\partial \beta_1} \\ \vdots \\ \frac{\partial S}{\partial \beta_I} \end{bmatrix}}$$

3) Part III: Let's use the same small example but now with $\mathbf{A} = \boldsymbol{X}^T\mathbf{X}$ ((2x2)-matrix). Note that this matrix $\boldsymbol{A}$ is symmetrical, i.e. $a_{21} = a_{12}$. Indeed:

$$\boldsymbol{X}^T\mathbf{X} = \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}$$

$$= \begin{bmatrix} {x_{11}}^2 + {x_{21}}^2 + {x_{31}}^2 & x_{12}x_{11} + x_{22}x_{21} + x_{32}x_{31} \\ x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32} & {x_{12}}^2 + {x_{22}}^2 + {x_{32}}^2 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Ordinary Least Squares Derived

Let’s differentiate it by parts:

$$\boxed{\frac{\partial S}{\partial \beta} = \begin{bmatrix} \frac{\partial S}{\partial \beta_1} \\ \vdots \\ \frac{\partial S}{\partial \beta_I} \end{bmatrix}}$$

$$S = \sum_{j=1}^{J} {e_j}^2 = \underbrace{\boldsymbol{y}^T\boldsymbol{y}}_{\text{Part I}} - \underbrace{\boldsymbol{2\hat{\beta}^T X^T y}}_{\text{Part II}} + \underbrace{\boldsymbol{\hat{\beta}^T X^T X \hat{\beta}}}_{\text{Part III}}$$

3) Part III:

$$\boldsymbol{\hat{\beta}^T X^T X \hat{\beta}} = \boldsymbol{\hat{\beta}^T A \hat{\beta}} = [\beta_1 \quad \beta_2] * \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$= [\beta_1 \quad \beta_2] * \begin{bmatrix} a_{11}\beta_1 + a_{12}\beta_2 \\ a_{21}\beta_1 + a_{22}\beta_2 \end{bmatrix}$$

$$= \beta_1 * (a_{11}\beta_1 + a_{12}\beta_2) + \beta_2 * (a_{21}\beta_1 + a_{22}\beta_2)$$

$$= a_{11}{\beta_1}^2 + 2a_{21}\beta_1\beta_2 + a_{22}{\beta_2}^2$$

Derivative with respect to β_1: $2a_{11}\beta_1 + 2a_{21}\beta_2 \overset{\boldsymbol{A}\text{ is symmetrical}}{=} 2a_{11}\beta_1 + 2a_{12}\beta_2$

Derivative with respect to β_2: $2a_{22}\beta_2 + 2a_{21}\beta_1 = 2a_{21}\beta_1 + 2a_{22}\beta_2$

➡ Derivative $= 2 * \begin{bmatrix} a_{11}\beta_1 + a_{12}\beta_2 \\ a_{21}\beta_1 + a_{22}\beta_2 \end{bmatrix} = 2\boldsymbol{A\beta} = 2\boldsymbol{X^T X \beta}$

Ordinary Least Squares Derived

Let’s differentiate it by parts:

$$S = \sum_{j=1}^{J} e_j{}^2 = \underbrace{\boldsymbol{y}^T\boldsymbol{y}}_{\text{Part I}} - \underbrace{\mathbf{2}\widehat{\boldsymbol{\beta}}^T\boldsymbol{X}^T\mathbf{y}}_{\text{Part II}} + \underbrace{\widehat{\boldsymbol{\beta}}^T\boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}}}_{\text{Part III}}$$

$$\boxed{\frac{\partial S}{\partial \beta} = \begin{bmatrix} \frac{\partial S}{\partial \beta_1} \\ \vdots \\ \frac{\partial S}{\partial \beta_I} \end{bmatrix}}$$

The total Derivative is thus: $\mathbf{2}\boldsymbol{X}^T\boldsymbol{X} - \mathbf{2}\boldsymbol{X}^T\mathbf{y}$

To find the minimum we have to set the derivative to zero:

$$\mathbf{2}\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} - \mathbf{2}\boldsymbol{X}^T\mathbf{y} = \mathbf{0}$$

$$\mathbf{2}\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} = \mathbf{2}\boldsymbol{X}^T\mathbf{y}$$

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\mathbf{y}$$

Note: taking the inverse is only possible if there are no two linearly dependent columns in the matrix. Hence, you have to take care not to put for instance the same regressor twice.

CHANGE OF TOPIC!

Good time for a break! ☺

Convolution and Linear Time Invariant Systems

Suppose you have a system.

input

System

output

The system is like a black box to you. You don’t know anything about it except that when you put something in, something comes out.

Convolution and Linear Time Invariant Systems

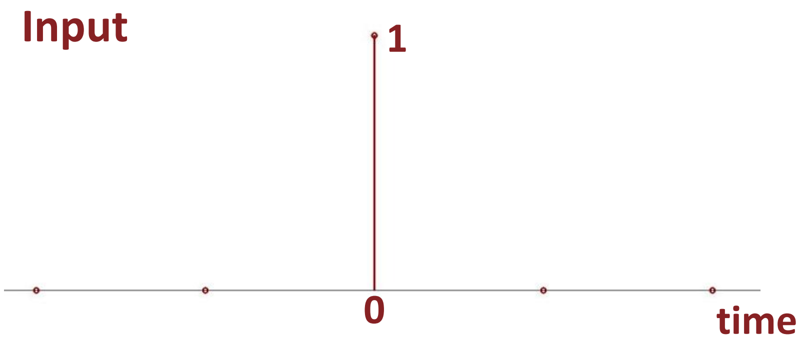
Suppose you want to understand this system.

For that purpose you put in the simplest signal you can imagine…

Convolution and Linear Time Invariant Systems

Suppose you want to understand this system.

For that purpose you put in the simplest signal you can imagine…

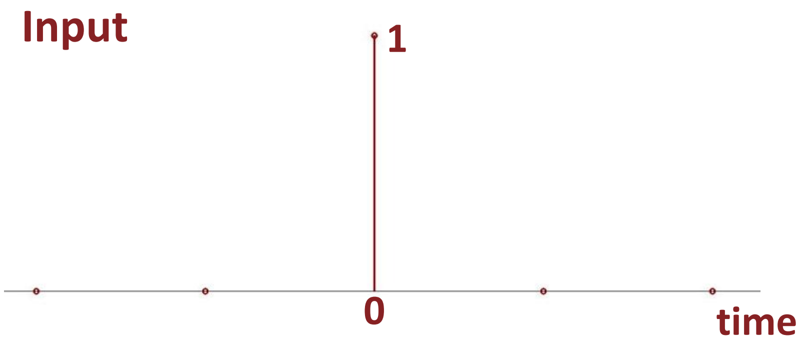


Convolution and Linear Time Invariant Systems

Suppose you want to understand this system.

For that purpose you put in the simplest signal you can imagine…

… and wait to see how the system reacts to it.

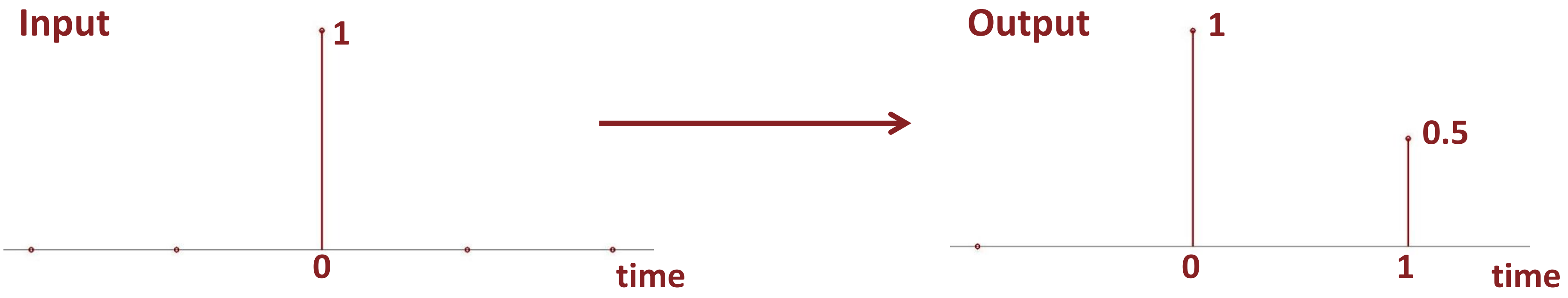


Convolution and Linear Time Invariant Systems

Suppose you want to understand this system.

For that purpose you put in the simplest signal you can imagine…

… and wait to see how the system reacts to it.

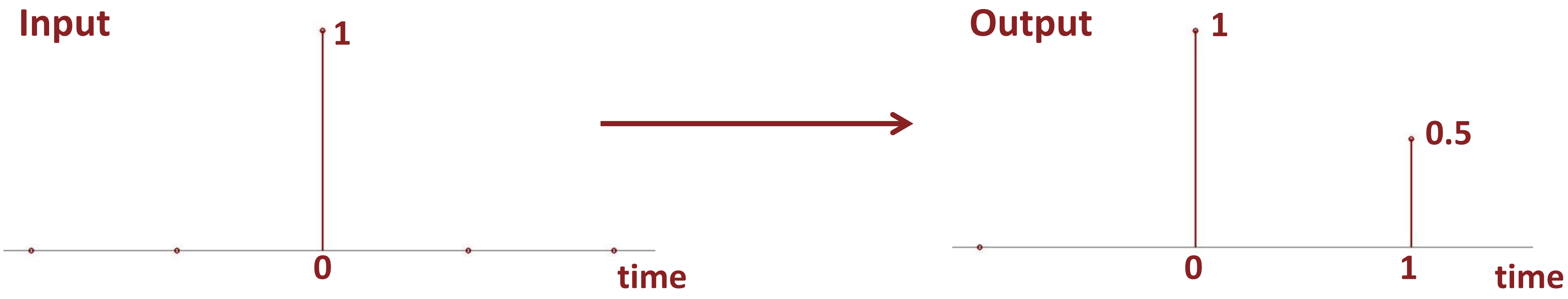


Convolution and Linear Time Invariant Systems

Suppose you want to understand this system.

For that purpose you put in the simplest signal you can imagine…

… and wait to see how the system reacts to it.



You do this many times and realize that the system always reacts to this simple signal in the same way.

The simple input as displayed here is called **impulse**. How the system reacts to it, whatever its form (the output displayed here was just an example), is called **impulse response function**.

Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Here we think about signals as discretized sequences of values. This means you only have discrete time points ($n = \cdots, -1,0,1,2,3 \ldots$) in which your signal assumes some value (e.g., the output signal of the previous slide assumes 1 at $n = 0$ and 0.5 at $n = 1$).

A reason for considering just discrete sequences is that computers actually also only deal with discrete sequences and not with continuous functions.

Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Here we think about signals as discretized sequences of values. This means you only have discrete time points ($n = \cdots, -1, 0, 1, 2, 3 \ldots$) in which your signal assumes some value (e.g., the output signal of the previous slide assumes 1 at $n = 0$ and 0.5 at $n = 1$).

A reason for considering just discrete sequences is that computers actually also only deal with discrete sequences and not with continuous functions.

The **impulse function** is defined as assuming the value one at time point zero and the value zero at all the other time points:

$$\delta[n] = \begin{cases} 0, if\ n \neq 0 \\ 1, if\ n = 0 \end{cases}$$

Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Here we think about signals as discretized sequences of values. This means you only have discrete time points ($n = \cdots, -1{,}0, 1, 2, 3 \ldots$) in which your signal assumes some value (e.g., the output signal of the previous slide assumes 1 at $n = 0$ and 0.5 at $n = 1$).

A reason for considering just discrete sequences is that computers actually also only deal with discrete sequences and not with continuous functions.

The **impulse function** is defined as assuming the value one at time point zero and the value zero at all the other time points:

$$\delta[n] = \begin{cases} 0, if\ n \neq 0 \\ 1, if\ n = 0 \end{cases}$$

You can shift it by p units to the left…

$$x[n] = \delta[n + p] = \begin{cases} 0, if\ n \neq 0 \\ 1, if\ n = -p \end{cases}$$

$$x[-p] = \delta[-p + p] = \delta[0] = 1$$

Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Here we think about signals as discretized sequences of values. This means you only have discrete time points ($n = \cdots, -1, 0, 1, 2, 3 \ldots$) in which your signal assumes some value (e.g., the output signal of the previous slide assumes 1 at $n = 0$ and 0.5 at $n = 1$).

A reason for considering just discrete sequences is that computers actually also only deal with discrete sequences and not with continuous functions.

The **impulse function** is defined as assuming the value one at time point zero and the value zero at all the other time points:

$$\delta[n] = \begin{cases} 0, if\ n \neq 0 \\ 1, if\ n = 0 \end{cases}$$

You can shift it by p units to the left…

$$x[n] = \delta[n + p] = \begin{cases} 0, if\ n \neq 0 \\ 1, if\ n = -p \end{cases}$$

$$x[-p] = \delta[-p + p] = \delta[0] = 1$$

… and to the right:

$$x[n] = \delta[n - p] = \begin{cases} 0, if\ n \neq 0 \\ 1, if\ n = p \end{cases}$$

$$x[p] = \delta[p - p] = \delta[0] = 1$$

Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Actually, the good thing about the impulse function is that any random sequence can be written as a linear combination of shifted and scaled versions of the impulse function.

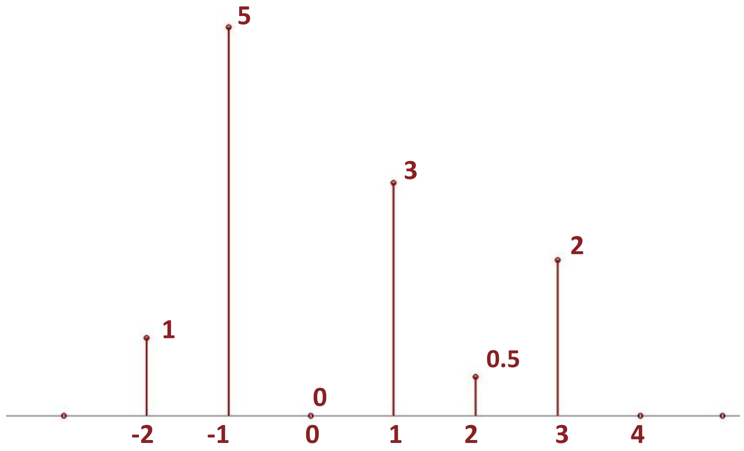
Imagine a complete random sequence $x[n]$:

Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Actually, the good thing about the impulse function is that any random sequence can be written as a linear combination of shifted and scaled versions of the impulse function.

Imagine a complete random sequence $x[n]$:

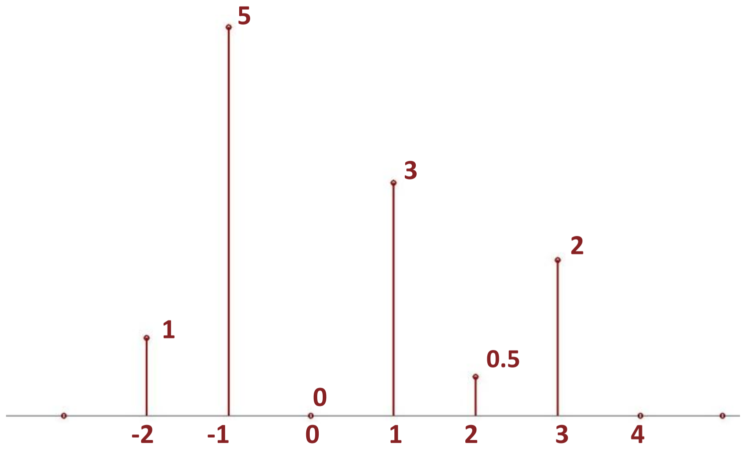


Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Actually, the good thing about the impulse function is that any random sequence can be written as a linear combination of shifted and scaled versions of the impulse function.

Imagine a complete random sequence $x[n]$:



You can write it as:

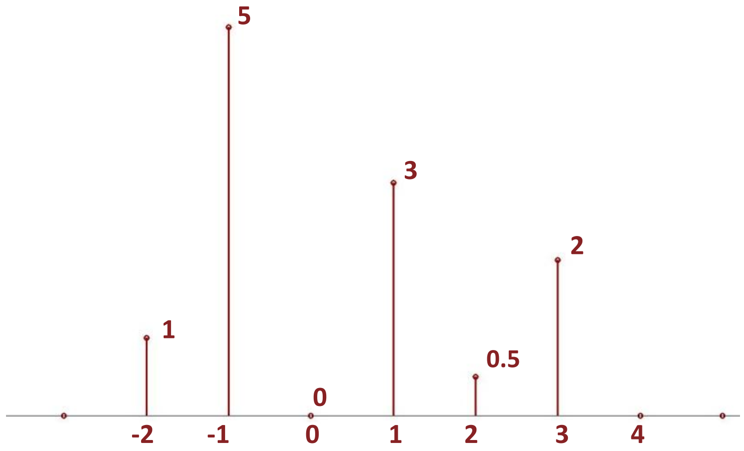
$$x[n] = 1 * \delta[n+2] + 5 * \delta[n+1] + 0 * \delta[n] + 3 * \delta[n-1] + 0.5 * \delta[n-2] + 2 * \delta[n-3]$$

Convolution and Linear Time Invariant Systems

A note about the impulse function and signals in general:

Actually, the good thing about the impulse function is that any random sequence can be written as a linear combination of shifted and scaled versions of the impulse function.

Imagine a complete random sequence $x[n]$:



So if we know how the system reacts to the impulse, to shifted and scaled versions of it and to additions of those, we might be able to know how the system reacts to all possible sequence you may put in!!

You can write it as:

$$x[n] = 1 * \delta[n+2] + 5 * \delta[n+1] + 0 * \delta[n] + 3 * \delta[n-1] + 0.5 * \delta[n-2] + 2 * \delta[n-3]$$

Convolution and Linear Time Invariant Systems

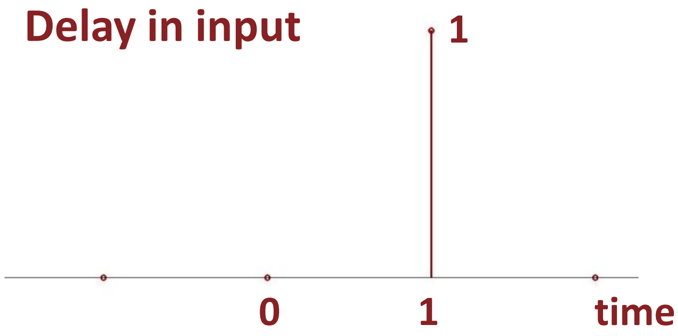
In our example we already know how the signal reacts to the impulse function. Let’s see how it reacts to a shifted (in this case delayed) version of it.

You proceed as before, you put in a shifted version and see how the system reacts. You repeat this several times and see that:

Convolution and Linear Time Invariant Systems

In our example we already know how the signal reacts to the impulse function. Let's see how it reacts to a shifted (in this case delayed) version of it.

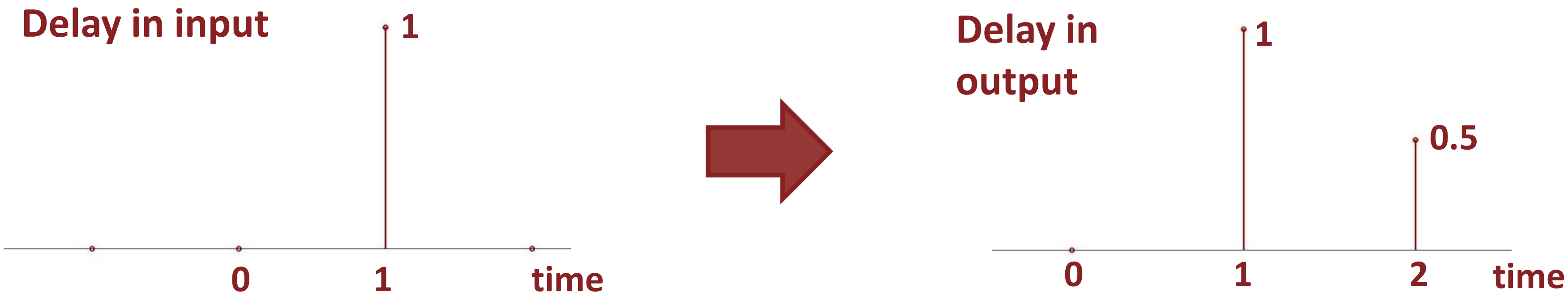
You proceed as before, you put in a shifted version and see how the system reacts. You repeat this several times and see that:



Convolution and Linear Time Invariant Systems

In our example we already know how the signal reacts to the impulse function. Let’s see how it reacts to a shifted (in this case delayed) version of it.

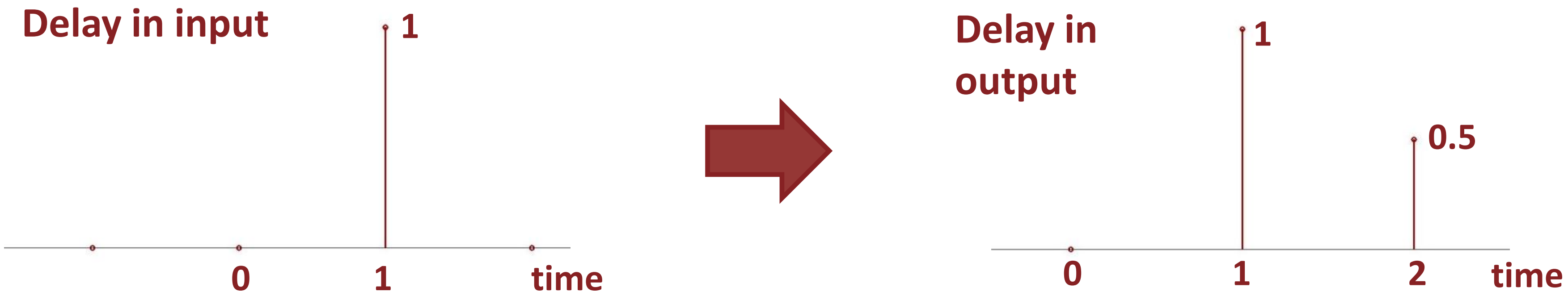
You proceed as before, you put in a shifted version and see how the system reacts. You repeat this several times and see that:



Convolution and Linear Time Invariant Systems

In our example we already know how the signal reacts to the impulse function. Let's see how it reacts to a shifted (in this case delayed) version of it.

You proceed as before, you put in a shifted version and see how the system reacts. You repeat this several times and see that:



You delayed the input by one and the output was also delayed by one. You do this for different shifts and see that the system always shifts the output by the same amount of the input shift. The system is said to be **time-invariant:** a shift in the input causes a corresponding shift in the output.

Convolution and Linear Time Invariant Systems

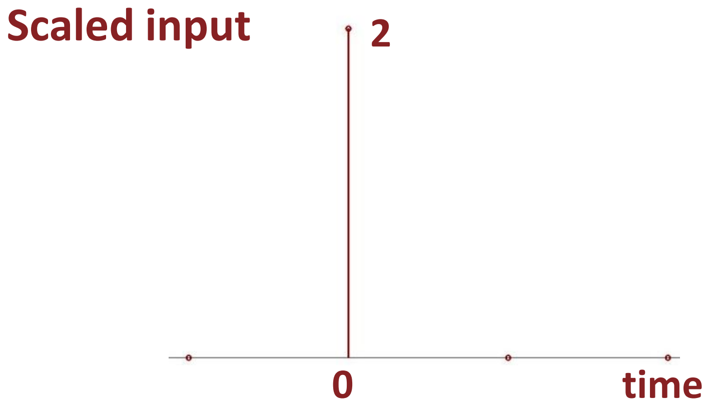
Let’s see how it reacts to a scaled version of it.

You put in a scaled version and see how the system reacts. You repeat this several times and with different scalings and see that:

Convolution and Linear Time Invariant Systems

Let’s see how it reacts to a scaled version of it.

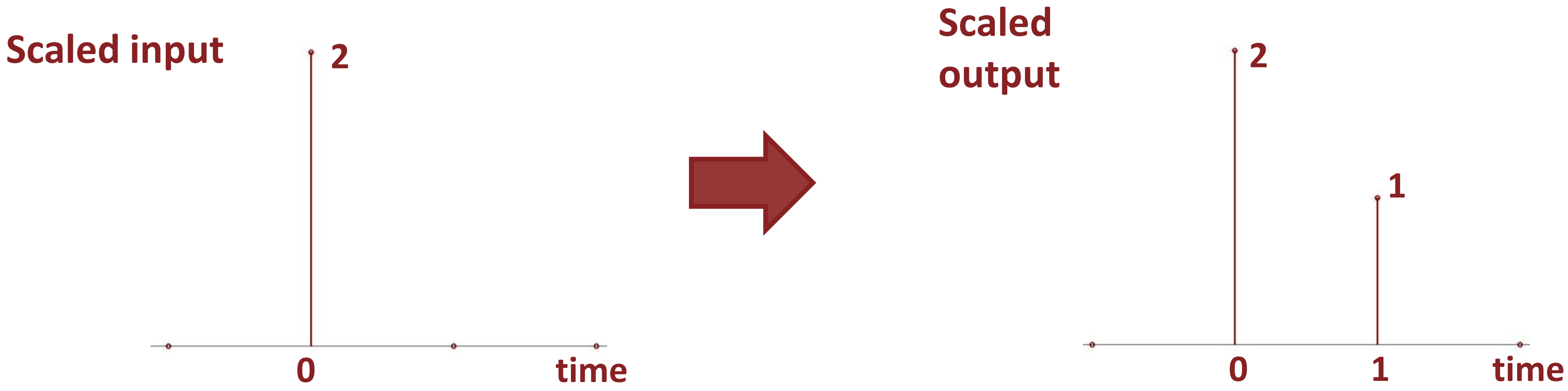
You put in a scaled version and see how the system reacts. You repeat this several times and with different scalings and see that:



Convolution and Linear Time Invariant Systems

Let’s see how it reacts to a scaled version of it.

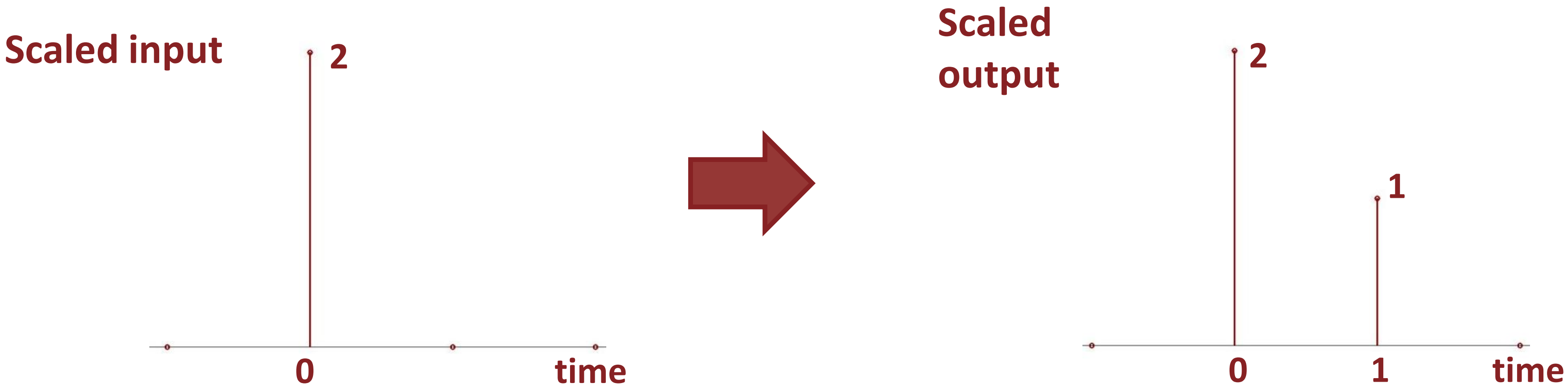
You put in a scaled version and see how the system reacts. You repeat this several times and with different scalings and see that:



Convolution and Linear Time Invariant Systems

Let’s see how it reacts to a scaled version of it.

You put in a scaled version and see how the system reacts. You repeat this several times and with different scalings and see that:

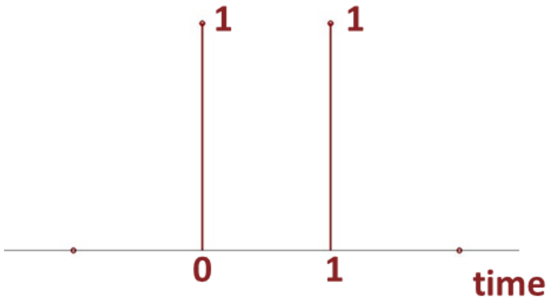


The system satisfies the scaling property:

If $x_1[n]$ has output $y_1[n]$, then $x[n] = a*x_1[n]$ has output $y[n] = a*y_1[n]$

Convolution and Linear Time Invariant Systems

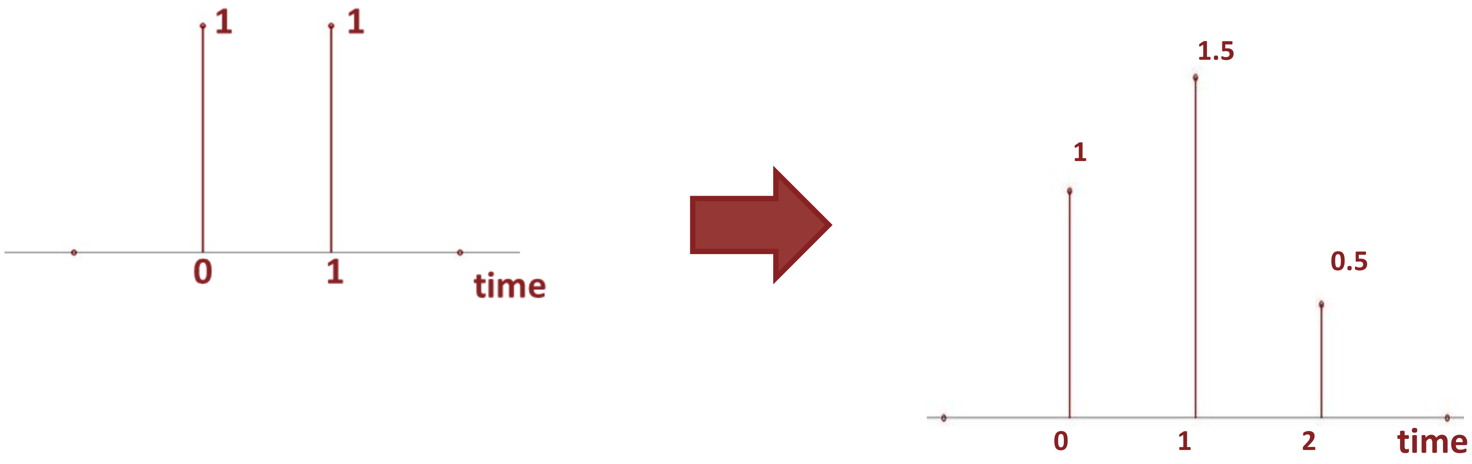
Now, let’s see how it reacts to a sum. To keep it simple, let’s consider the sum of the impulse and a shifted version of itself, for instance:



Convolution and Linear Time Invariant Systems

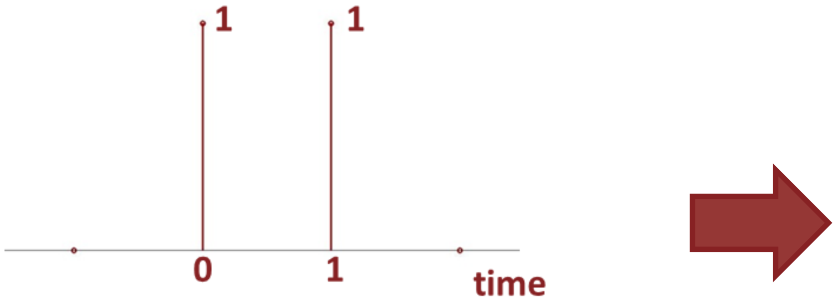
Now, let’s see how it reacts to a sum. To keep it simple, let’s consider the sum of the impulse and a shifted version of itself, for instance:

And you see the output is:

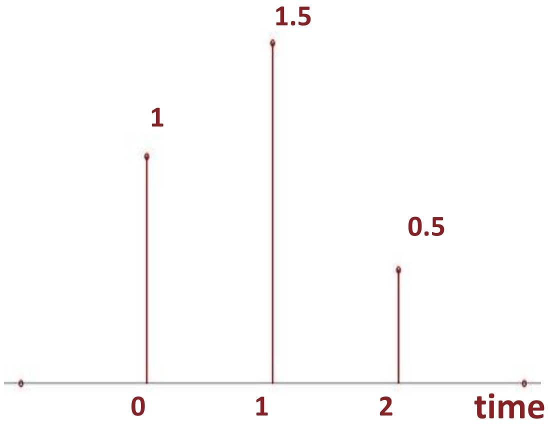


Convolution and Linear Time Invariant Systems

Now, let’s see how it reacts to a sum. To keep it simple, let’s consider the sum of the impulse and a shifted version of itself, for instance:



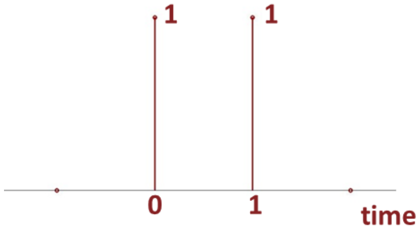
And you see the output is:

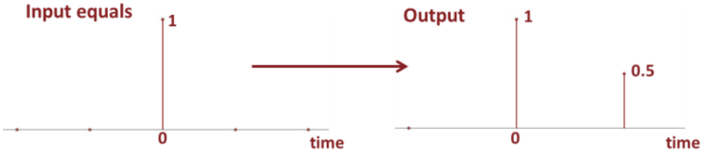


To understand better how this output came into place, let’s consider the separate contributions of each individual “stick” at any time point.

Convolution and Linear Time Invariant Systems

Input:



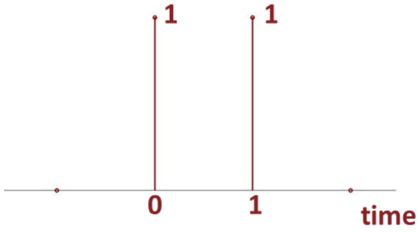


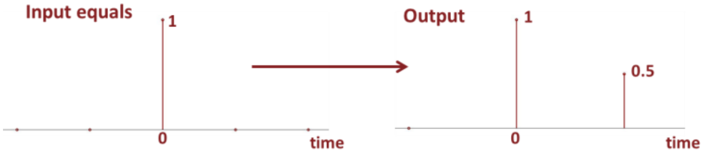
Contribution 1st impulse

Contribution 2nd impulse

Convolution and Linear Time Invariant Systems

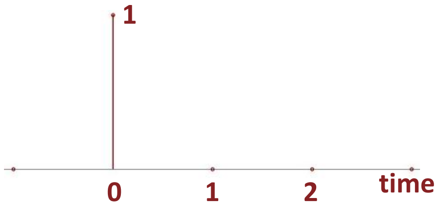
Input:





time 0

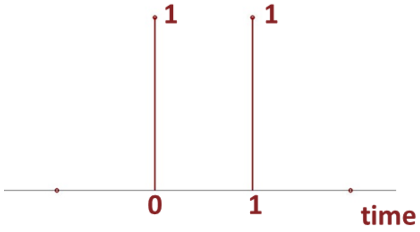
Contribution 1st impulse

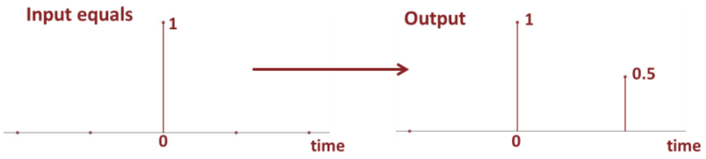


Contribution 2nd impulse

Convolution and Linear Time Invariant Systems

Input:



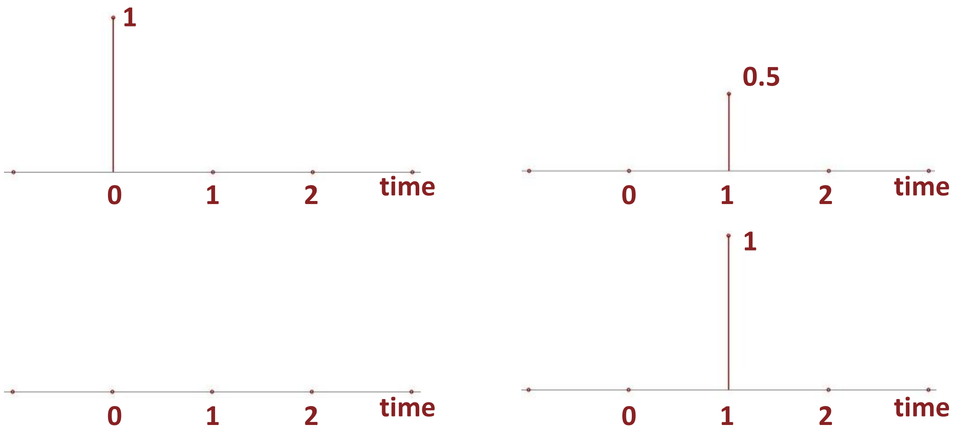


time 0

time 1

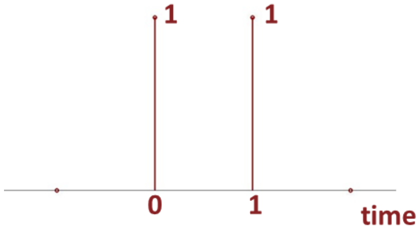
Contribution 1st impulse

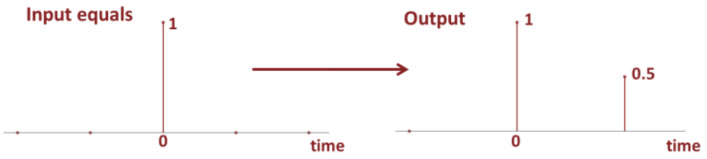
Contribution 2nd impulse



Convolution and Linear Time Invariant Systems

Input:





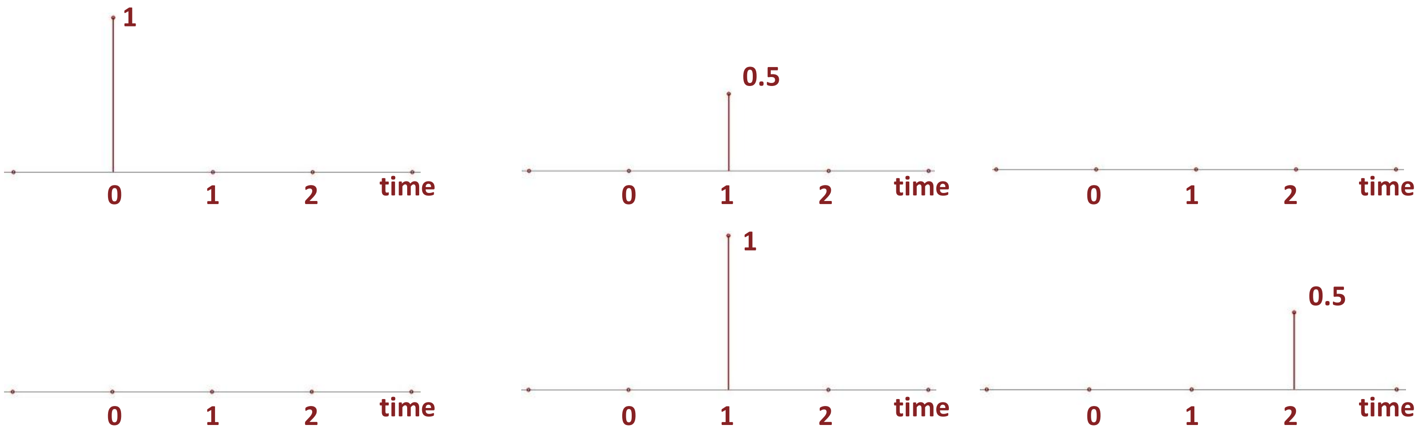
time 0

time 1

time 2

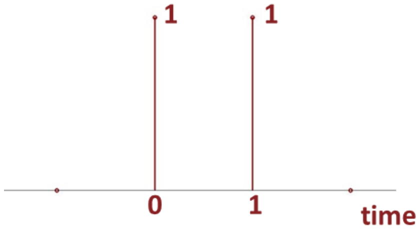
Contribution 1^{st} impulse

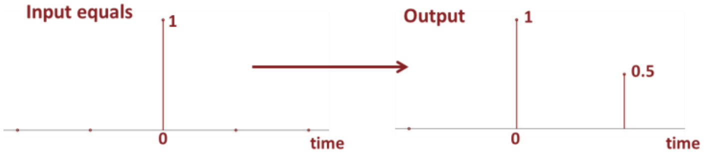
Contribution 2^{nd} impulse



Convolution and Linear Time Invariant Systems

Input:





time 0

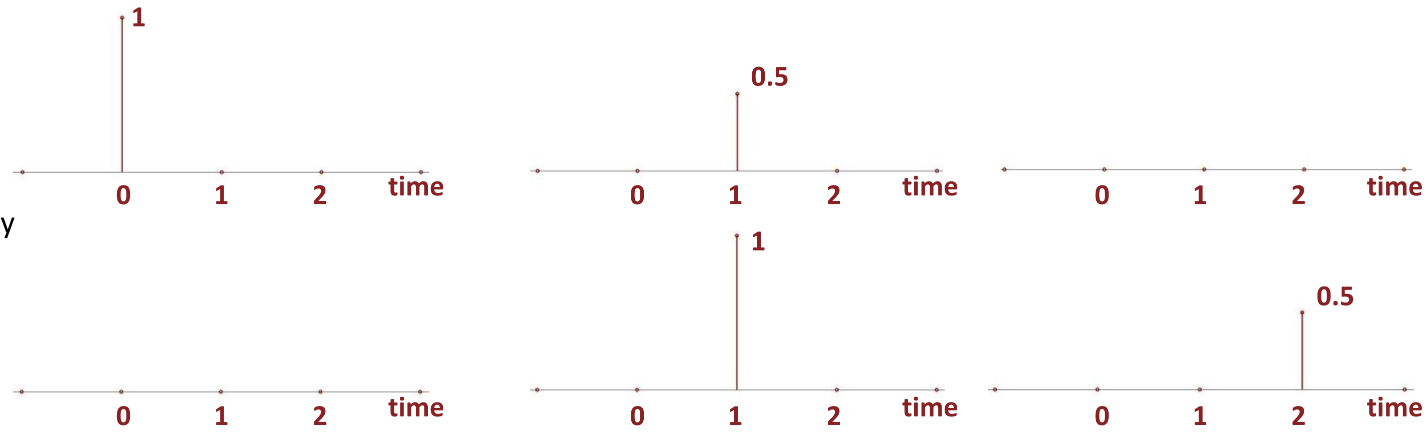
time 1

time 2

Contribution 1st impulse

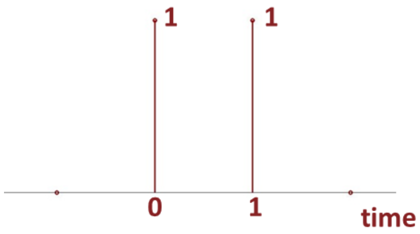
+ add entry by entry

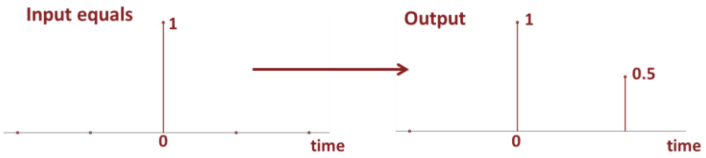
Contribution 2nd impulse



Convolution and Linear Time Invariant Systems

Input:





time 0

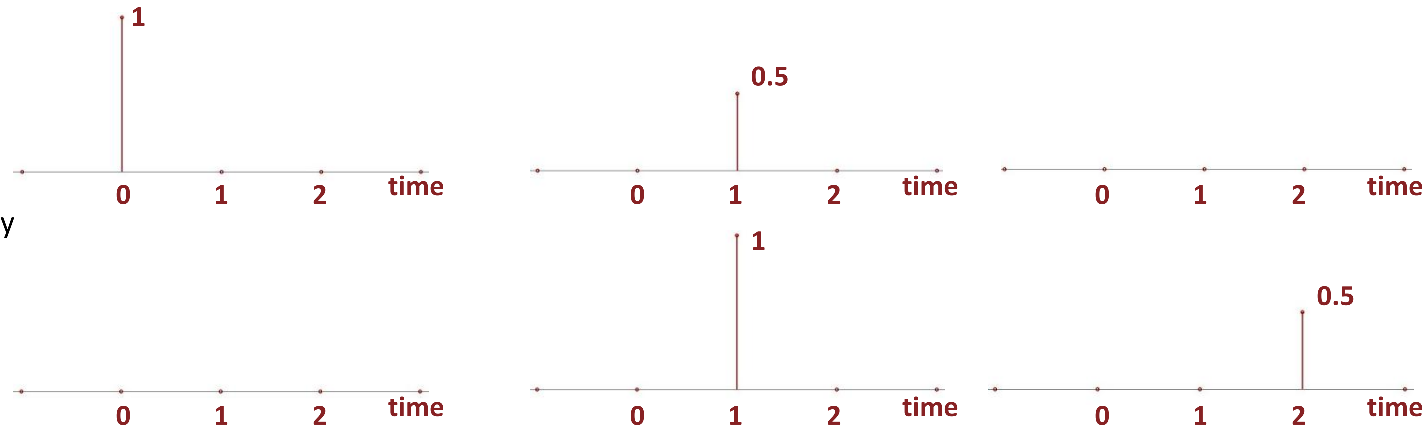
time 1

time 2

Contribution 1^{st} impulse

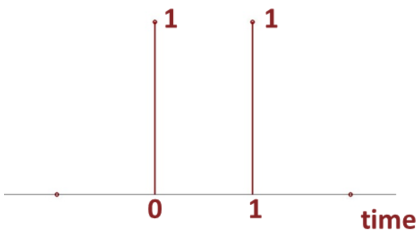
⊕ add entry by entry

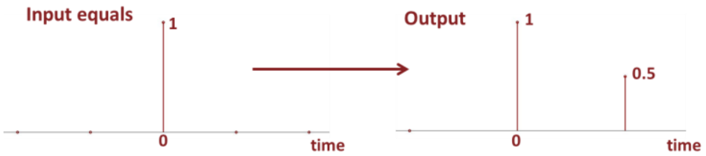
Contribution 2^{nd} impulse



Convolution and Linear Time Invariant Systems

Input:





time 0

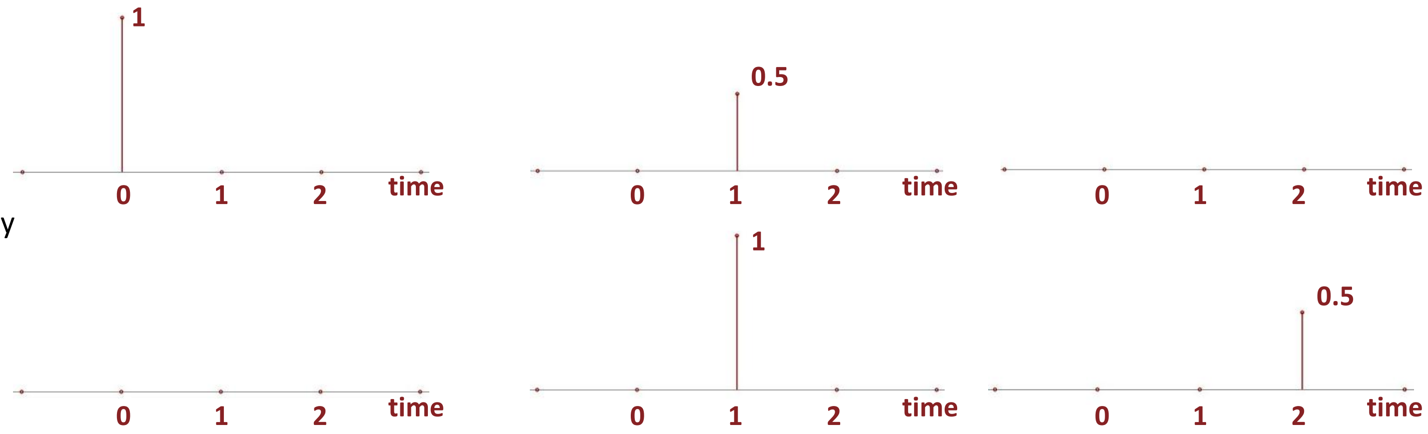
time 1

time 2

Contribution 1^{st} impulse

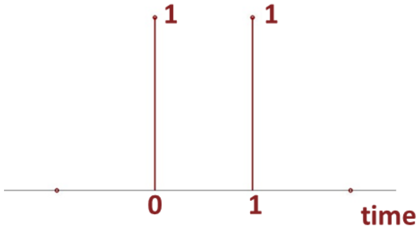
add entry by entry

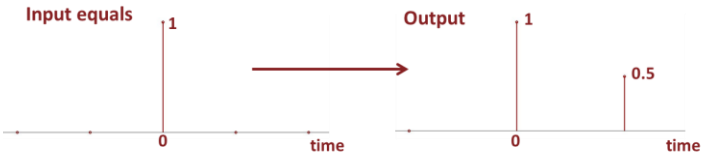
Contribution 2^{nd} impulse



Convolution and Linear Time Invariant Systems

Input:

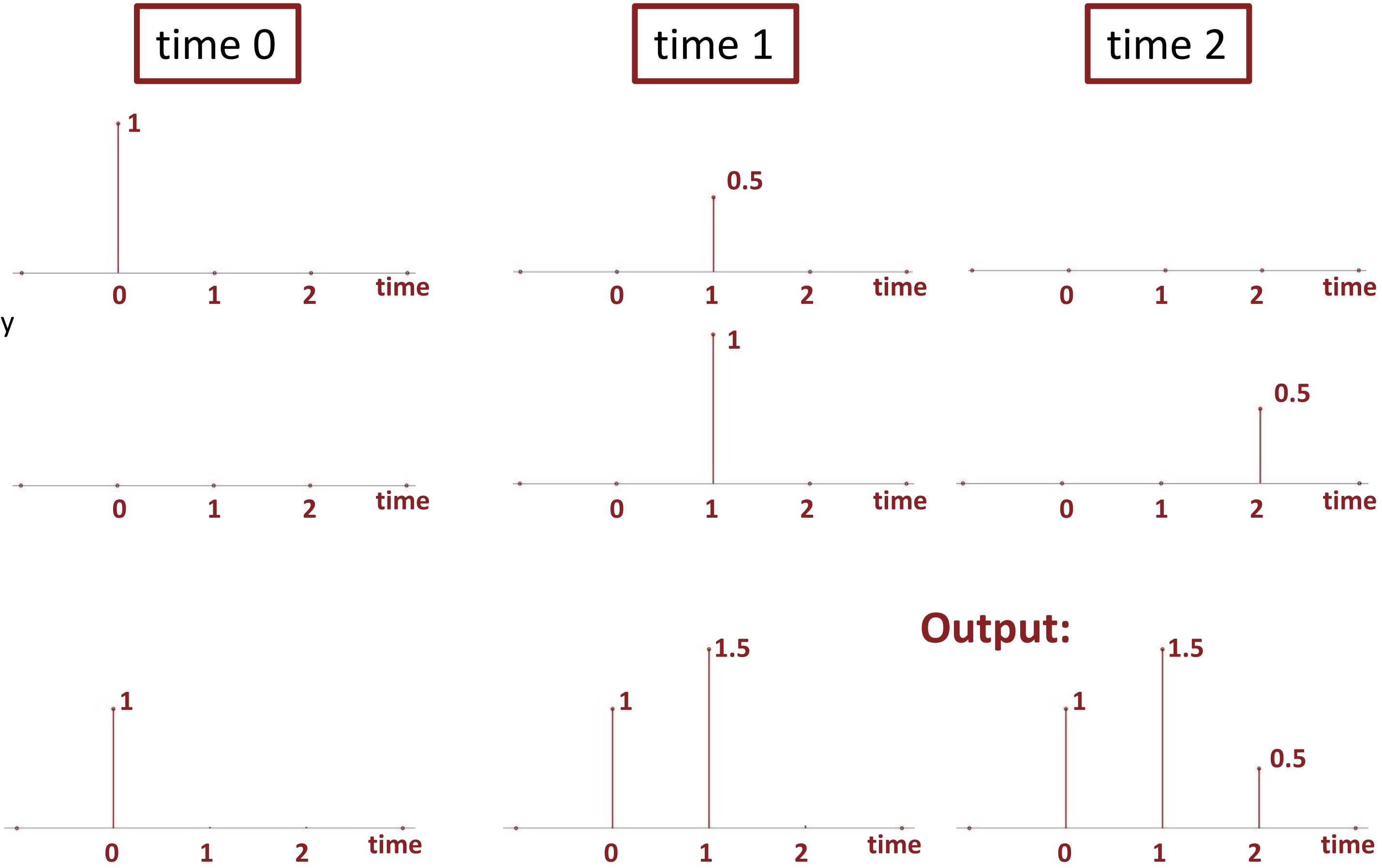




Contribution 1^{st} impulse

+ add entry by entry

Contribution 2^{nd} impulse



Convolution and Linear Time Invariant Systems

The output is actually just the sum of the contributions of the individual sticks at each time point.

Hence, the system satisfies the additivity property:

If $x_1[n]$ has output $y_1[n]$ and $x_2[n]$ has output $y_2[n]$,

then $x[n] \; = \; (x_1[n] + x_2[n])$ has output $y[n] \; = \; (y_1[n] \; + \; y_2[n])$

Convolution and Linear Time Invariant Systems

The output is actually just the sum of the contributions of the individual sticks at each time point.

Hence, the system satisfies the additivity property:

If $x_1[n]$ has output $y_1[n]$ and $x_2[n]$ has output $y_2[n]$,

then $x[n] = (x_1[n] + x_2[n])$ has output $y[n] = (y_1[n] + y_2[n])$

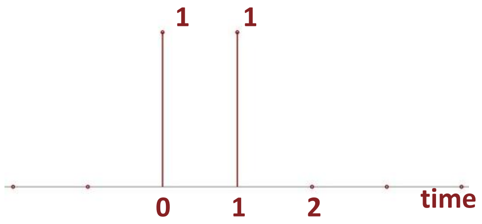
When a system satisfies the additivity and the scaling properties, it is said to be **linear**.

Linear and *time-invariant* systems are just systems that satisfy these two properties simultaneously. These two properties together allow for the system to be characterized by its impulse response function. If you know the impulse response function of a system you can calculate the output to any input by *convolving* your input with the impulse response function. The convolution is just another, more (computationally) practical, way of looking at the sum we saw in the previous slide.

Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.

Input



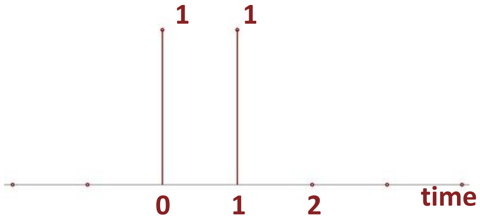
Flipped Impulse Response Function

Output

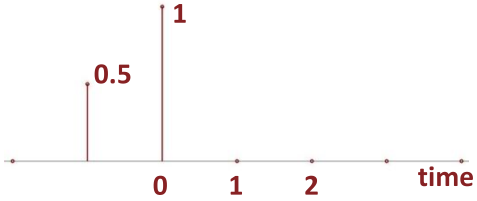
Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.

Input



Flipped Impulse Response Function

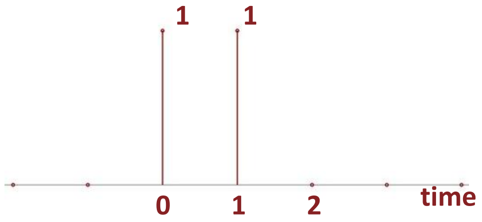


Output

Convolution and Linear Time Invariant Systems

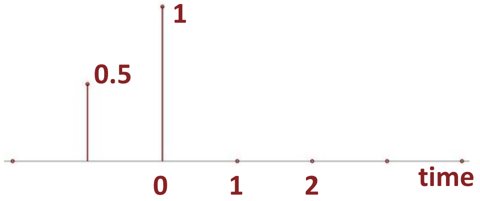
Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.

Input



multiply overlapping entries and add the results ⊕

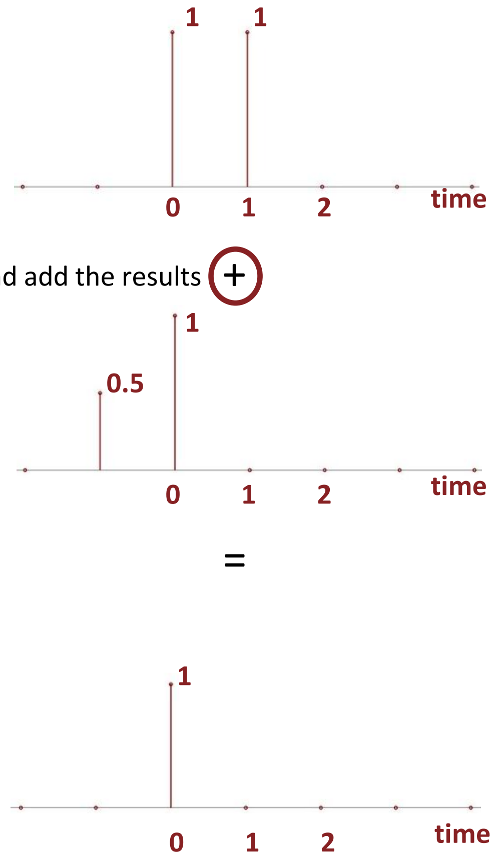
Flipped Impulse Response Function



Output

Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.



Input

multiply overlapping entries and add the results

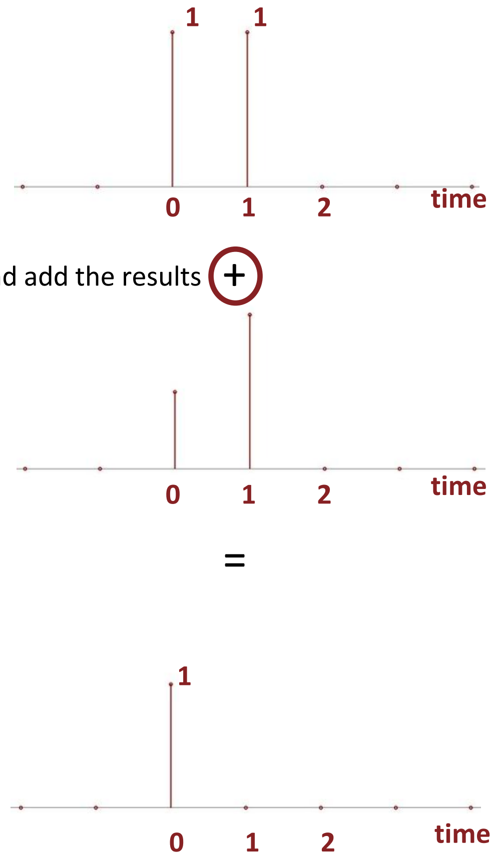
Flipped Impulse Response Function

Output

At time point zero only the first impulse contributes to the output

Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.



Input

multiply overlapping entries and add the results +

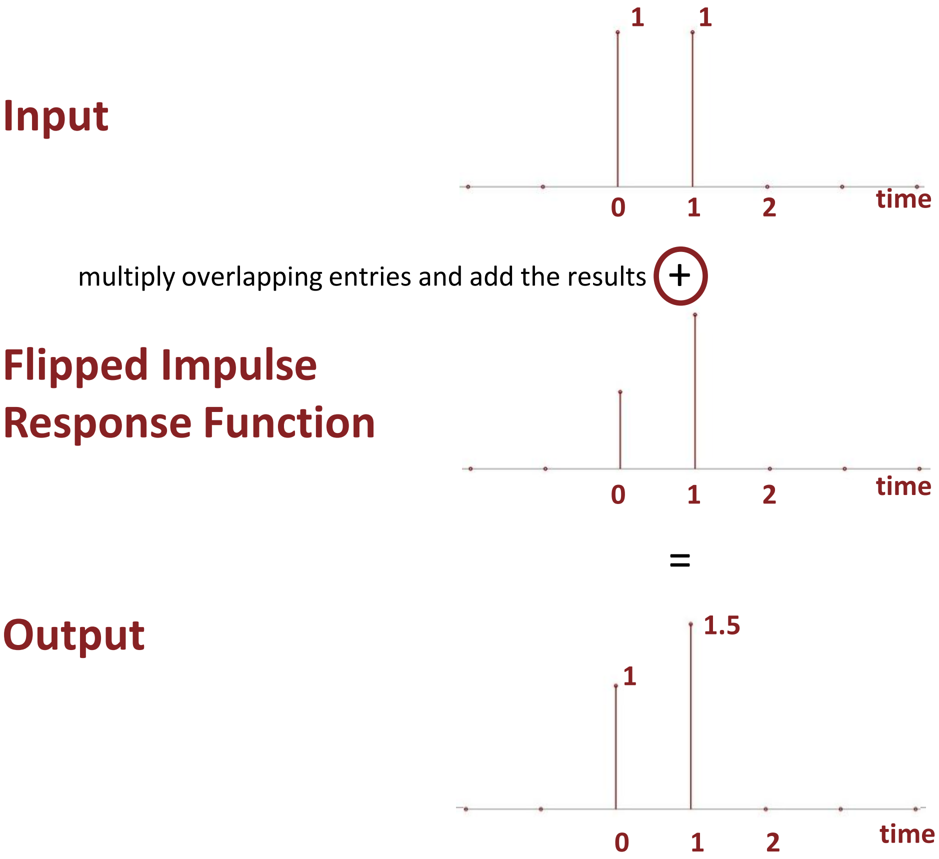
Flipped Impulse Response Function

=

Output

Convolution and Linear Time Invariant Systems

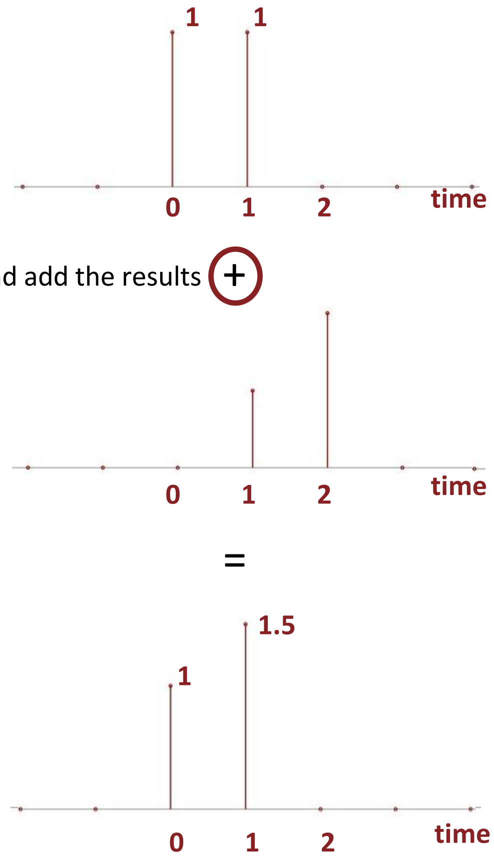
Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.



At time point one the first impulse contributes with 0.5 while the second impulse has just started to kick in and contributes with 1 to the output

Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.



Input

multiply overlapping entries and add the results ⊕

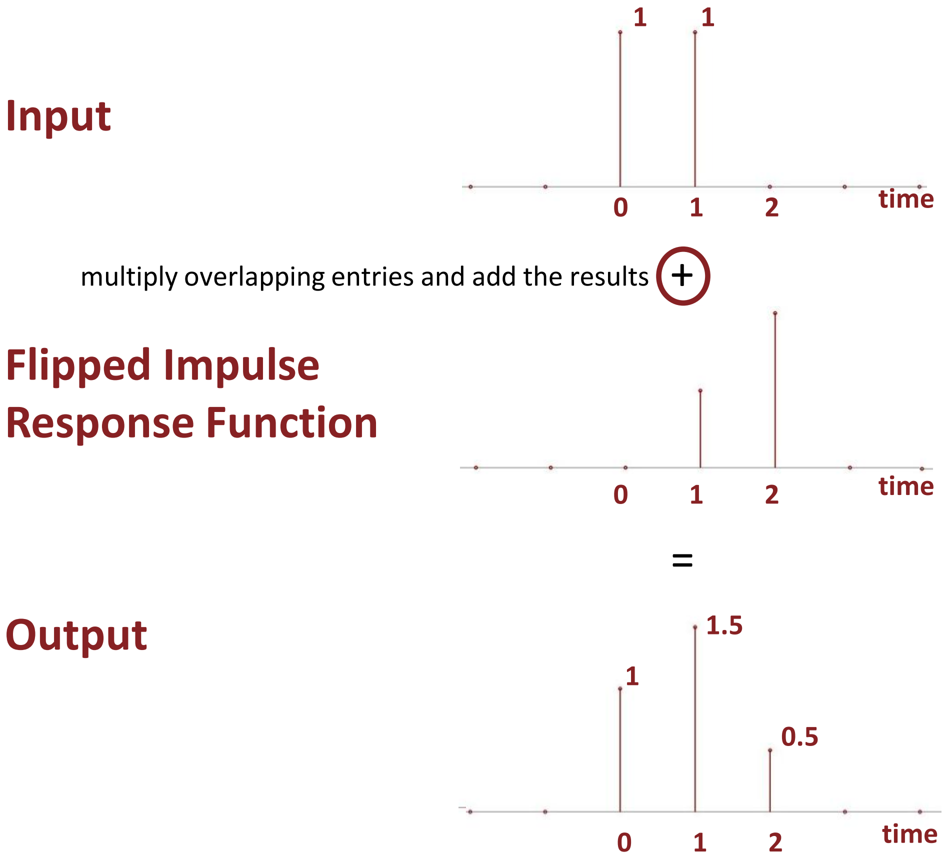
Flipped Impulse Response Function

=

Output

Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.



Input

multiply overlapping entries and add the results ⊕

Flipped Impulse Response Function

=

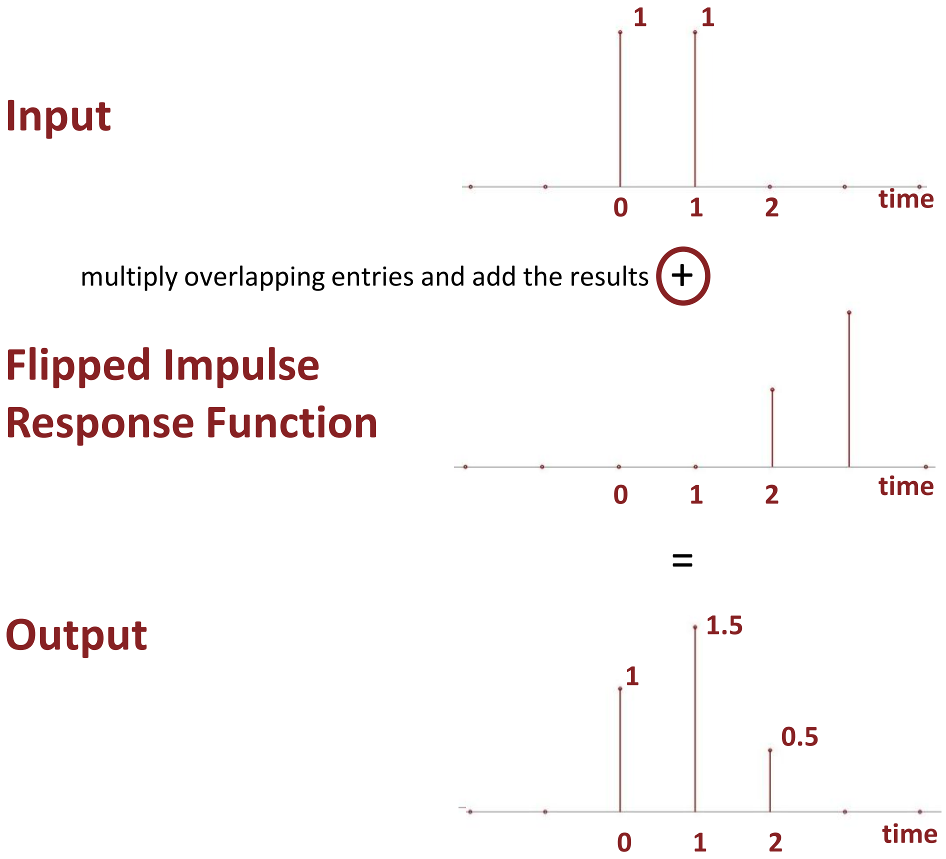
Output

At time point two the first impulse is already quiet while the second impulse contributes with 0.5 to the output

Convolution and Linear Time Invariant Systems

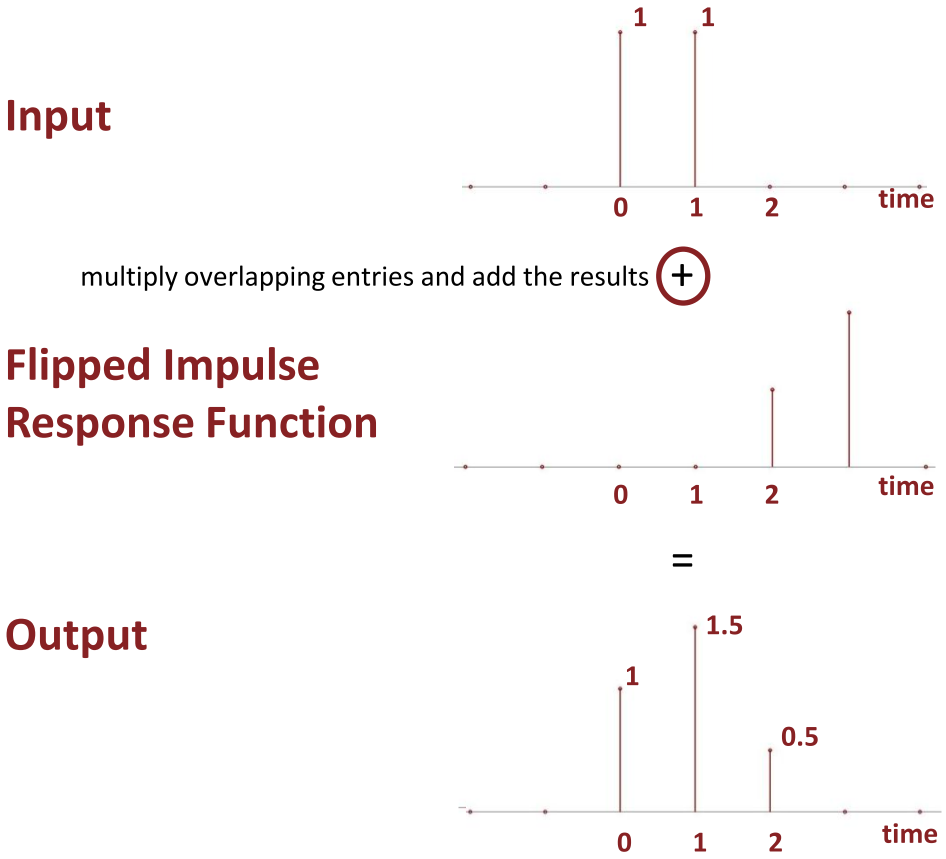
Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.

After that all is quiet.



Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.



Input

multiply overlapping entries and add the results ⊕

Flipped Impulse Response Function

=

Output

After that all is quiet.

CONVOLUTION

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k].$$

$$f \otimes g(t) = \int_{0}^{t} f(\tau)g(t-\tau)d\tau$$

Convolution and Linear Time Invariant Systems

Indeed, looking at the individual responses and then adding them is equivalent to convolution. The convolution can be visually thought of as follows.

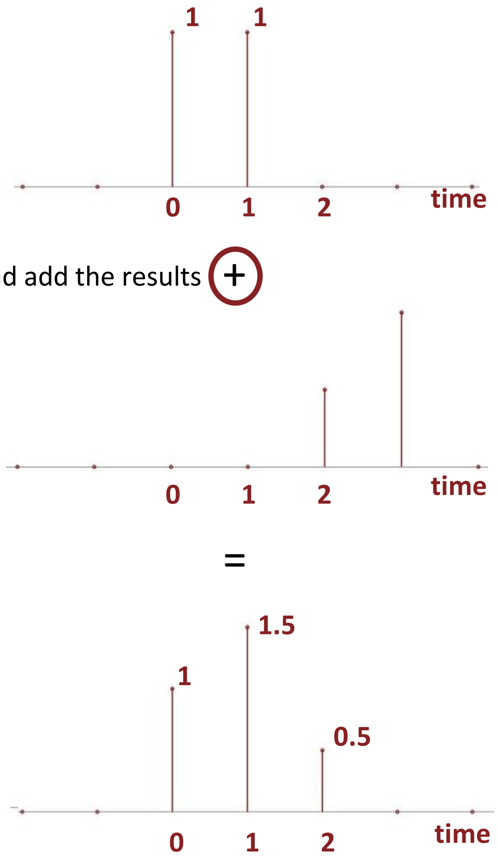
After that all is quiet.

Input

multiply overlapping entries and add the results

Flipped Impulse Response Function

Output



CONVOLUTION

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k].$$

$$f \otimes g(t) = \int_{0}^{t} f(\tau)g(t-\tau)d\tau$$

A LTI is completely characterized by its impulse response function!

LTIs & SPM?

SPM assumes the brain behaves like a linear time-invariant system.

The impulse corresponds to an event (e.g. brief flash of light). If the events don't come too close in time, this assumption has proven to work rather well.

LTIs & SPM?

SPM assumes the brain behaves like a linear time-invariant system.

The impulse corresponds to an event (e.g. brief flash of light). If the events don’t come too close in time, this assumption has proven to work rather well.

What is the impulse response function of the brain?

LTIs & SPM?

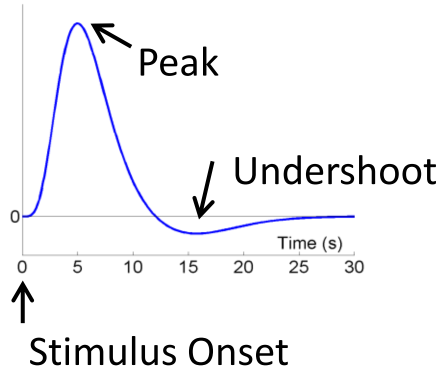
SPM assumes the brain behaves like a linear time-invariant system.

The impulse corresponds to an event (e.g. brief flash of light). If the events don't come too close in time, this assumption has proven to work rather well.

What is the impulse response function of the brain?

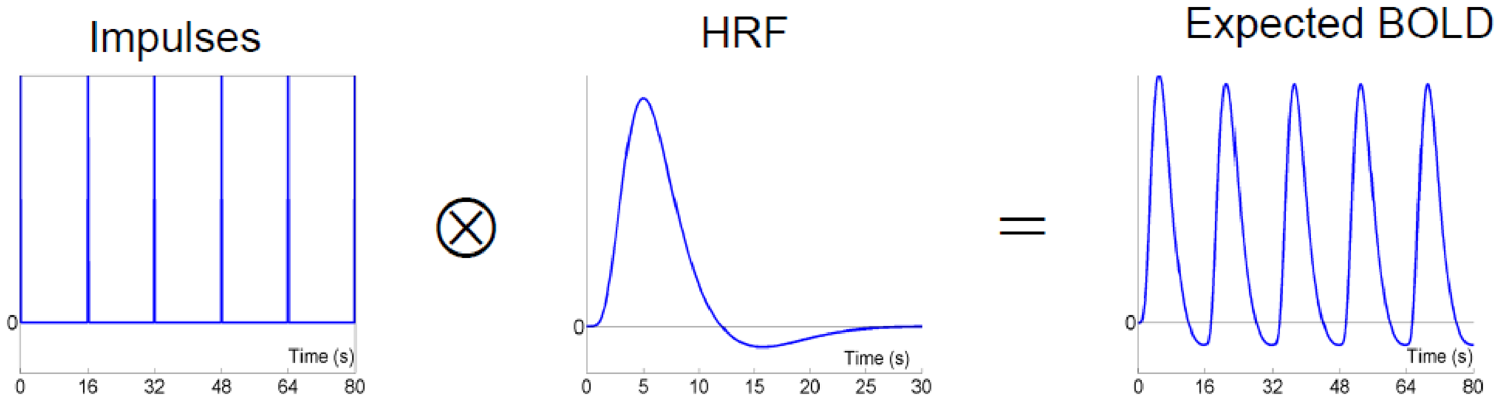
The canonical HRF is a "typical" BOLD impulse response function and is characterised by (the subtraction of) two gamma functions.

Haemodynamic Response Function



LTIs & SPM?

So, in order to know how the brain reacts to a certain sequence of events, you have to convolve your regressors with the HRF



LTIs & SPM?

Limits of the HRF

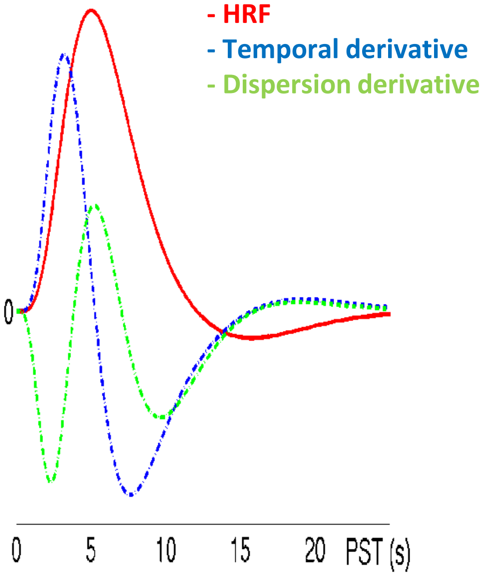
- General shape of the BOLD impulse response similar across early sensory regions, such as V1 and S1

- Variability across higher cortical regions, presumably due mainly to variations in the vasculature of different regions

- Considerable variability across people

These types of variability can be accommodated by expanding the HRF in terms of temporal basis functions.

- **Temporal derivative**: captures differences in latency of peak response
- **Dispersion derivative:** captures differences in duration of peak response

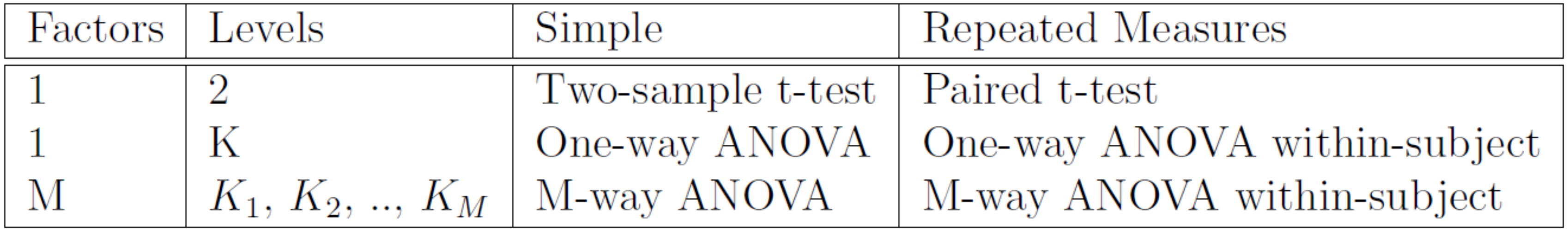


CHANGE OF TOPIC!

Good time for a break! ☺

Specifying Design Matrices

Different types of ANOVAs:



Factors	Levels	Simple	Repeated Measures
1	2	Two-sample t-test	Paired t-test
1	K	One-way ANOVA	One-way ANOVA within-subject
M	K_1, K_2, ..., K_M	M-way ANOVA	M-way ANOVA within-subject

As the ANOVA is a generalization of the two-sample t-test for more than two groups, the repeated measures ANOVA is the generalization of the paired t-test for more than two groups.
Here the design matrices for the ANOVAs will not be presented, but they can be obtained from those of the two-sample and paired t-tests by extending the number of groups.

One sample t-test:

Does the group (we have just one group in this case) have any significant activation?

Suppose Y_{1j} is a group of random variables and assume that $Y_{1j} \overset{iid}{\sim} \mathcal{N}(\mu_1, \sigma^2)$.
For each member j of the group, you can write:

$$Y_{1j} = \mu_1 + \varepsilon_{1j} \qquad \text{where} \qquad \varepsilon_{1j} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

Design matrix for 8 participants:

$$\boldsymbol{y} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\boldsymbol{X}} * \beta_1 + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = \beta_1 + \varepsilon_{11} \\ Y_{12} = \beta_1 + \varepsilon_{12} \\ Y_{13} = \beta_1 + \varepsilon_{13} \\ Y_{14} = \beta_1 + \varepsilon_{14} \\ Y_{15} = \beta_1 + \varepsilon_{15} \\ Y_{16} = \beta_1 + \varepsilon_{16} \\ Y_{17} = \beta_1 + \varepsilon_{17} \\ Y_{18} = \beta_1 + \varepsilon_{18} \end{cases} \Rightarrow$$

Specifying Design Matrices

One sample t-test:

Does the group (we have just one group in this case) have any significant activation?

Suppose Y_{1j} is a group of random variables and assume that $Y_{1j} \overset{iid}{\sim} \mathcal{N}(\mu_1, \sigma^2)$.
For each member j of the group, you can write:

$$Y_{1j} = \mu_1 + \varepsilon_{1j} \qquad \text{where} \qquad \varepsilon_{1j} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

Design matrix for 8 participants:

$$\boldsymbol{y} = \underbrace{\begin{bmatrix} \boxed{\begin{matrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{matrix}} \end{bmatrix}}_{\boldsymbol{X}} * \beta_1 + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = \beta_1 + \varepsilon_{11} \\ Y_{12} = \beta_1 + \varepsilon_{12} \\ Y_{13} = \beta_1 + \varepsilon_{13} \\ Y_{14} = \beta_1 + \varepsilon_{14} \\ Y_{15} = \beta_1 + \varepsilon_{15} \\ Y_{16} = \beta_1 + \varepsilon_{16} \\ Y_{17} = \beta_1 + \varepsilon_{17} \\ Y_{18} = \beta_1 + \varepsilon_{18} \end{cases} \Rightarrow \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}}$$

Specifying Design Matrices

One sample t-test:

Does the group (we have just one group in this case) have any significant activation?

Suppose Y_{1j} is a group of random variables and assume that $Y_{1j} \overset{iid}{\sim} \mathcal{N}(\mu_1, \sigma^2)$.
For each member j of the group, you can write:

$$Y_{1j} = \mu_1 + \varepsilon_{1j} \quad \text{where} \quad \varepsilon_{1j} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

Design matrix for 8 participants:

$$\boldsymbol{y} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\boldsymbol{X}} * \beta_1 + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = \beta_1 + \varepsilon_{11} \\ Y_{12} = \beta_1 + \varepsilon_{12} \\ Y_{13} = \beta_1 + \varepsilon_{13} \\ Y_{14} = \beta_1 + \varepsilon_{14} \\ Y_{15} = \beta_1 + \varepsilon_{15} \\ Y_{16} = \beta_1 + \varepsilon_{16} \\ Y_{17} = \beta_1 + \varepsilon_{17} \\ Y_{18} = \beta_1 + \varepsilon_{18} \end{cases} \overset{\boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}}}{\Longrightarrow} \boxed{\beta_1 = \mu_1}$$

Specifying Design Matrices

Two sample t-test:

You have two groups. Do these two groups have significant difference in brain activation?

Suppose Y_{1j} and Y_{2j} are two independent groups of random variables and assume that $Y_{qj} \overset{iid}{\sim} \mathcal{N}(\mu_q, \sigma^2)$, for $q = 1,2$. The null hypothesis is $\mathcal{H}: \mu_1 = \mu_2$.

For all members of group 1, you can write:

$$Y_{1j} = \mu_1 + \varepsilon_{1j} \quad \text{where} \quad \varepsilon_{1j} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

And for members of group 2, you can write:

$$Y_{2j} = \mu_2 + \varepsilon_{2j} \quad \text{where} \quad \varepsilon_{2j} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model I:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases} \quad \Longrightarrow$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model I:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases} \Longrightarrow \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model I:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases} \overset{\boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}}}{\Longrightarrow} \boxed{\beta_1 = \mu_1}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model I:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

$$\begin{matrix} \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \\ \Longrightarrow \\ \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}} \end{matrix} \qquad \boxed{\beta_1 = \mu_1}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model I:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

$$\begin{matrix} \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \\ \Longrightarrow \\ \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}} \end{matrix} \qquad \begin{matrix} \boxed{\beta_1 = \mu_1} \\ \boxed{\beta_2 = \mu_2} \end{matrix}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model I:

contrasts

$$(1,0).\beta = \mu_1$$
$$(0,1).\beta = \mu_2$$
$$(1,-1).\beta = \mu_1 - \mu_2$$
$$(0.5,0.5).\beta = mean(\mu_1, \mu_2)$$

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

$$\boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}}$$

$$\Longrightarrow$$

$$\boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}}$$

$$\boxed{\beta_1 = \mu_1}$$

$$\boxed{\beta_2 = \mu_2}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases} \Longrightarrow$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

$$
\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases} \Longrightarrow \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}}
$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

$$\boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \Longrightarrow \quad \boxed{\beta_1 + \beta_2 = \mu_1}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases} \Longrightarrow \begin{matrix} \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \\ \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}} \end{matrix} \quad \boxed{\beta_1 + \beta_2 = \mu_1}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

$$\begin{matrix} \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \\ \Longrightarrow \\ \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}} \end{matrix} \qquad \begin{matrix} \boxed{\beta_1 + \beta_2 = \mu_1} \\ \boxed{\beta_2 = \mu_2} \end{matrix}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

$$\boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \quad \Longrightarrow \quad \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}}$$

$$\boxed{\beta_1 + \beta_2 = \mu_1}$$

$$\boxed{\beta_2 = \mu_2}$$

$$\Downarrow$$

$$\beta_1 = \mu_1 - \mu_2$$

$$\beta_2 = \mu_2$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

contrasts

$(1{,}1).\beta = \mu_1$

$(0{,}1).\beta = \mu_2$

$(1{,}0).\beta = \mu_1 - \mu_2$

$(0.5{,}1).\beta = mean(\mu_1, \mu_2)$

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

$$\Longrightarrow \begin{array}{c} \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \\ \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}} \end{array}$$

$$\boxed{\beta_1 + \beta_2 = \mu_1}$$

$$\boxed{\beta_2 = \mu_2}$$

$$\Downarrow$$

$$\beta_1 = \mu_1 - \mu_2$$

$$\beta_2 = \mu_2$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model II:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 1 * \beta_2 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + \varepsilon_{26} \end{cases}$$

contrasts

$(1,1).\beta = \mu_1$

$(0,1).\beta = \mu_2$

$(1,0).\beta = \mu_1 - \mu_2$

$(0.5,1).\beta = mean(\mu_1, \mu_2)$

Example of calculation of contrast:

$$\begin{aligned} \mu_1 - \mu_2 &= \beta_1 + \beta_2 - \beta_2 \\ &= \beta_1 \end{aligned}$$

Hence, the contrast for the difference has only a weight on β_1 and equals c=(1,0)

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model III:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{26} \end{cases} \Longrightarrow$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model III:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{26} \end{cases} \Longrightarrow \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model III:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{26} \end{cases}$$

$$\boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \quad \Longrightarrow \quad \boxed{\beta_1 + \beta_3 = \mu_1}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model III:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{26} \end{cases}$$

$$\boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \quad \Longrightarrow \quad \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}}$$

$$\boxed{\beta_1 + \beta_3 = \mu_1}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model III:

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{11} \\ Y_{12} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{12} \\ Y_{13} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{13} \\ Y_{14} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{14} \\ Y_{15} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{15} \\ Y_{16} = 1 * \beta_1 + 0 * \beta_2 + 1 * \beta_3 + \varepsilon_{16} \\ Y_{21} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{21} \\ Y_{22} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{22} \\ Y_{23} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{23} \\ Y_{24} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{24} \\ Y_{25} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{25} \\ Y_{26} = 0 * \beta_1 + 1 * \beta_2 + 1 * \beta_3 + \varepsilon_{26} \end{cases}$$

$$\begin{array}{c} \boxed{Y_{1j} = \mu_1 + \varepsilon_{1j}} \\ \Longrightarrow \\ \boxed{Y_{2j} = \mu_2 + \varepsilon_{2j}} \end{array} \qquad \begin{array}{c} \boxed{\beta_1 + \beta_3 = \mu_1} \\ \\ \boxed{\beta_2 + \beta_3 = \mu_2} \end{array}$$

Specifying Design Matrices

Two sample t-test:

Assume you have two groups of 6 participants each. You can have three possible models.

Model III:

contrasts

$$(1,0,1).\beta = \mu_1$$
$$(0,1,1).\beta = \mu_2$$
$$(1,-1,0).\beta = \mu_1 - \mu_2$$
$$(0.5,0.5,1).\beta = mean(\mu_1,\mu_2)$$

$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \boldsymbol{\varepsilon} \Leftrightarrow \begin{cases} Y_{11} = 1*\beta_1 + 0*\beta_2 + 1*\beta_3 + \varepsilon_{11} \\ Y_{12} = 1*\beta_1 + 0*\beta_2 + 1*\beta_3 + \varepsilon_{12} \\ Y_{13} = 1*\beta_1 + 0*\beta_2 + 1*\beta_3 + \varepsilon_{13} \\ Y_{14} = 1*\beta_1 + 0*\beta_2 + 1*\beta_3 + \varepsilon_{14} \\ Y_{15} = 1*\beta_1 + 0*\beta_2 + 1*\beta_3 + \varepsilon_{15} \\ Y_{16} = 1*\beta_1 + 0*\beta_2 + 1*\beta_3 + \varepsilon_{16} \\ Y_{21} = 0*\beta_1 + 1*\beta_2 + 1*\beta_3 + \varepsilon_{21} \\ Y_{22} = 0*\beta_1 + 1*\beta_2 + 1*\beta_3 + \varepsilon_{22} \\ Y_{23} = 0*\beta_1 + 1*\beta_2 + 1*\beta_3 + \varepsilon_{23} \\ Y_{24} = 0*\beta_1 + 1*\beta_2 + 1*\beta_3 + \varepsilon_{24} \\ Y_{25} = 0*\beta_1 + 1*\beta_2 + 1*\beta_3 + \varepsilon_{25} \\ Y_{26} = 0*\beta_1 + 1*\beta_2 + 1*\beta_3 + \varepsilon_{26} \end{cases}$$

$$Y_{1j} = \mu_1 + \varepsilon_{1j}$$

$$\Longrightarrow$$

$$Y_{2j} = \mu_2 + \varepsilon_{2j}$$

$$\beta_1 + \beta_3 = \mu_1$$

$$\beta_2 + \beta_3 = \mu_2$$

Specifying Design Matrices

Paired t-test:

You have two groups in which observations in one group can be paired with observations in the other group

Use it when:

- Before-and-after observations on the same subjects
- Comparison of two different methods of measurement where the measurements are applied to the same subjects

If you have 2 measures per participant and N participants the k^{th} response ($k = 1,2$) from the n^{th} participant is modelled as:

$$y_{nk} = \tau_k + \pi_n + e_{nk}$$

where τ_k are the treatment effects, π_n the subject effects and e_{nk} the residual errors.

Specifying Design Matrices

Paired t-test:

Assume you perform two treatments on 6 different participants.

$$\boldsymbol{y} = \left[\underbrace{\begin{matrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{matrix}}_{\text{treatments}} \; \underbrace{\begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{matrix}}_{\text{participants}}\right] * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8 \end{bmatrix} + \boldsymbol{\varepsilon} \quad \overset{y_{nk} = \tau_k + \pi_n + e_{nk}}{\Longrightarrow} \quad \begin{cases} \beta_1 = \tau_1 \\ \beta_2 = \tau_2 \\ \beta_3 = \pi_1 \\ \beta_4 = \pi_2 \\ \beta_5 = \pi_3 \\ \beta_6 = \pi_4 \\ \beta_7 = \pi_5 \\ \beta_8 = \pi_6 \end{cases}$$

THE END