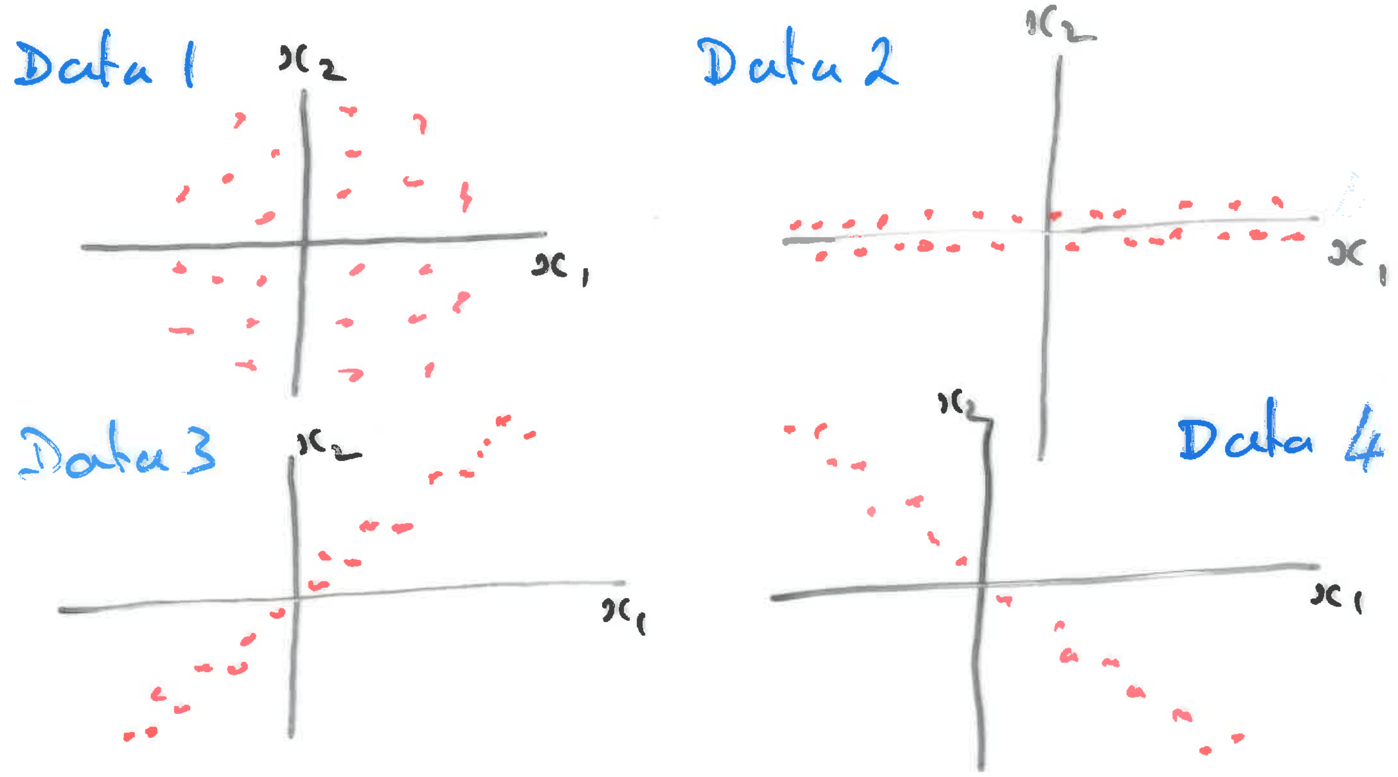
Recap $v_1, \dots, v_n \in \mathbb{R}^p$, $\bar{v} = \frac{1}{n}(v_1 + \dots + v_n) = \underline{0}$

$$v_k = \begin{pmatrix} x_{k1} \\ \vdots \\ x_{kp} \end{pmatrix} \qquad c_{ij} = \frac{1}{n} \sum_{k=1}^{n} x_{ki}\, x_{kj}$$

x_{*i} and x_{*j} uncorrelated if $c_{ij} = 0$

Covariance matrix $c = \begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & & \\ c_{p1} & \cdots & c_{pp} \end{pmatrix}$ is symmetric

Example $(p = 2)$



Data	c_{11}	c_{22}	$c_{12} = c_{21}$
1	large	large	0
2	large	small	0
3	large	large	large positive
4	large	large	large negative

In Data 2 we can project $\mathbb{R}^2 \to \mathbb{R}$ onto the x_1-axis without loosing much info about variation in the data.

In Data 3 and 4 we can rotate the data, and then project onto x_1-axis without loosing much info.

Principal Component Analysis (PCA) is a method for finding the relevant rotations. $\boxed{4}$

In general, the covariance matrix depends on $v_1, \dots, v_n \in \mathbb{R}^p$.

Say

$$C = C(v_1, \dots, v_n)$$

<u>Aim of PCA</u> we'd like to find a distance preserving homomorphism

$$\phi : \mathbb{R}^p \to \mathbb{R}^p, \quad v \mapsto \phi(v)$$

such that the covariance matrix

$$C = C(\phi v_1, \phi v_2, \dots, \phi v_n)$$

has $c_{ij} = 0$ for $i \neq j$.

Recall

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{R}^p, \quad \|v\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_p^2}$$

$v_1, v_2 \in \mathbb{R}^p,$

Euclidean distance $d(v_1, v_2) = \|v_1 - v_2\|$

Note

$$v^t v = \|v\|^2$$

for $\phi : \mathbb{R}^p \to \mathbb{R}^p,\ v \mapsto Av$ to be distance preserving (i.e. to be an isometry) we need

$$\|v\| = \|Av\|$$

i.e. we just need

$$v^t v = (Av)^t (Av)$$

or

$$v^t v = v^t A^t A v$$

i.e. we just need $A^t A = I$.

Defn A square matrix A is **orthogonal** if $A^t A = I$.

Aim of PCA

We'd like to find an orthogonal $p \times p$ matrix A such that the covariance matrix

$$C(Av_1, \dots, Av_n)$$

has $c_{ij} = 0$ for $i \neq j$.

Consider $p \times n$ matrix

$$X = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & \vdots \end{pmatrix}$$

Covariance matrix $C = \frac{1}{n} X X^t$

Consider

$$Y = \begin{pmatrix} | & | & & | \\ \phi v_1 & \phi v_2 & \cdots & \phi v_n \\ | & | & & | \end{pmatrix} = AX$$

In PCA we aim to find an orthogonal matrix A such that

$$\frac{1}{n} Y Y^t = \underbrace{\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{pmatrix}}_{\Lambda}$$

i.e. we want

$$\frac{1}{n} (A x)(A x)^t = \Lambda$$

i.e.

$$\frac{1}{n} A x x^t A^t = \Lambda$$

<u>Aim of PCA</u>

We'd like to find an orthogonal $p \times p$ matrix A such that, for $C = \frac{1}{n} x x^t$, the matrix

$$A C A^t = \Lambda \qquad (*)$$

is diagonal.

$$A^{-1} = A^t \quad \text{if} \quad A A^t = I$$

$$(*) \qquad C A^t = A^t \Lambda$$

If $A = \begin{pmatrix} — & r_1 & — \\ — & r_2 & — \\ & \vdots & \\ — & r_p & — \end{pmatrix}$ the $(*)$ means

$$C \begin{pmatrix} | \\ r_k \\ | \end{pmatrix} = \lambda_k \begin{pmatrix} | \\ r_k \\ | \end{pmatrix} \quad \text{for} \quad 1 \leq k \leq p.$$

<u>**Aim of PCA**</u>

We'd like to find an orthogonal $p \times p$ matrix A whose rows are (transposes of) eigenvectors of the covariance matrix.

Q Can we always perform PCA on any $v_1, \dots, v_n \in \mathbb{R}^p$?

Note: $C = XX^t$ is symmetric since $C^t = (XX^t)^t = XX^t = C$.

A Yes we can always do PCA because:

Spectral Theorem Let C be any real symmetric matrix. Then there exists an orthogonal matrix A whose columns are eigenvectors of C. This ensures

$$A^t C A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{pmatrix}$$

with $\lambda_1, \dots, \lambda_p$ the corresponding eigenvalues.