# 2D Random vector

$$\underset{(2\times1)}{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

simultaneous pdf

$f(x_1, x_2) =$ "$P(X_1 = x_2, X_2 = x_2)$"

complete information.

# Marginal pdf's

$$f(x_1) = \int f(x_1, x_2)\, dx_2$$

$$f(x_2) = \int f(x_1, x_2)\, dx_1$$

Mean $\mu$ vector $(2 \times 1)$ $= \begin{pmatrix} E[x_1] \\ E[x_2] \end{pmatrix}$

Marginal variance

$$\sigma_1^2 = \sigma_{11} = E[(x_1 - \mu_1)^2] \geq 0$$

$$\sigma_2^2 = \sigma_{22} = E[(x_2 - \mu_2)^2] \geq 0$$

# Covariance

$$\sigma_{12} = \mathrm{cov}(x_1, x_2)$$

$$= E\left[(x_1 - \mu_1)(x_2 - \mu_2)\right]$$

pos

0

neg

# Covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

Symmetric $\Rightarrow$ positive

eigenvalues $\lambda_1$, $\lambda_2$

$\underline{\lambda_1 \geq 0}$ , $\underline{\lambda_2 \geq 0}$

# Correlation coefficient

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \, \sigma_2}$$

correlation matrix

$$\underset{2\times 2}{\rho} = \begin{array}{cc} x_1 & \\ x_2 & \end{array} \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}$$

$$-1 \leq \rho_{12} \leq 1$$

# Generalized variance

$$|\Sigma| = \det(\Sigma) = \prod_{i=1}^{2} \lambda_i$$

(2x2)

Measure of amount of random variation.

Lower is less chaotic and thus results in more oder

$p$ D $\quad X = [x_1, x_2 \ldots x_p]^T$

$(p \times 1)$

<u>random</u>

<u>vector</u> mean vector $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

$(p \times 1)$

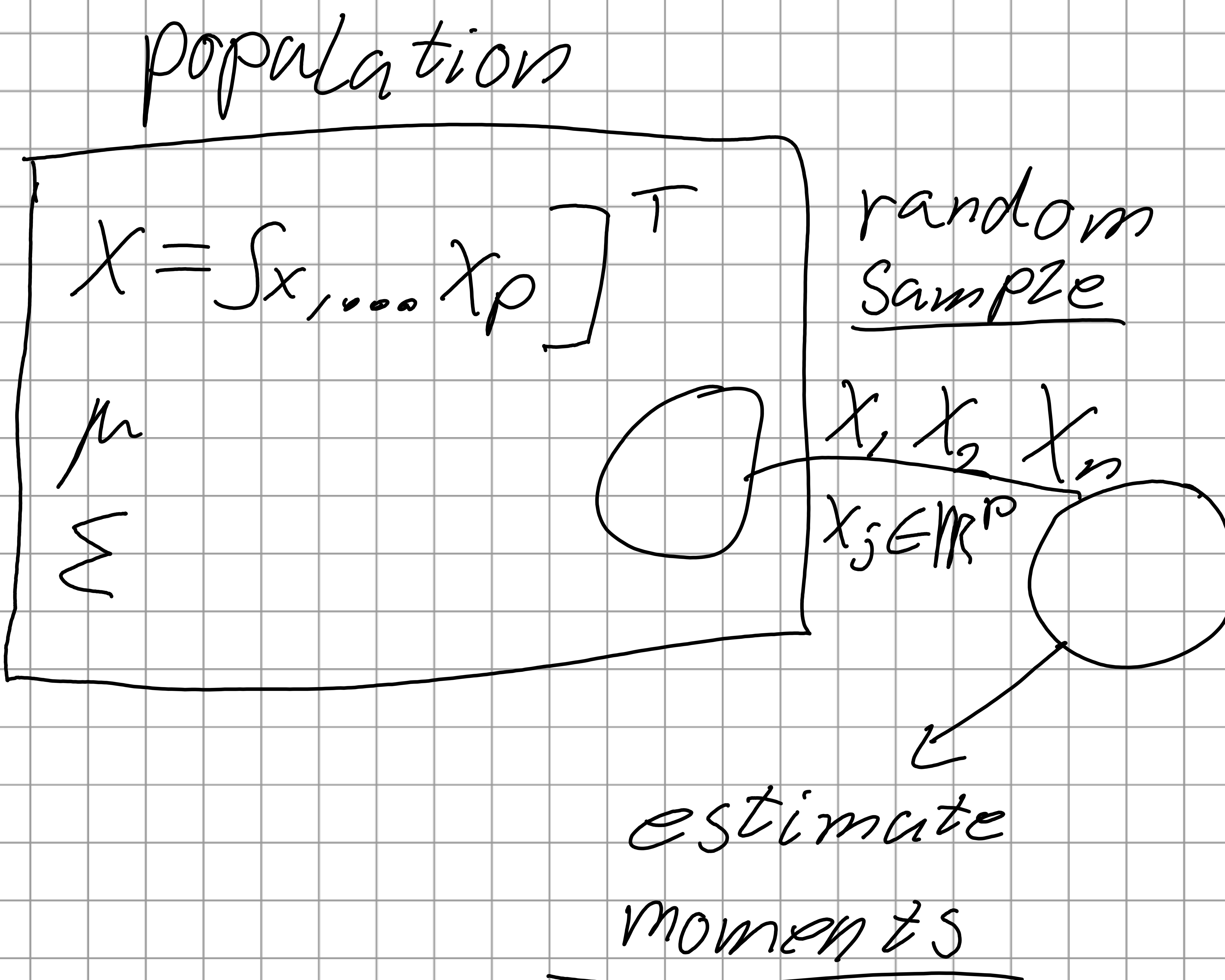
covar matrix

$$\underset{p \times p}{\Sigma} = \begin{array}{cc} \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_p^2 \end{bmatrix} & \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_p \end{array} \\ x_1 \; x_2 \ldots x_p & \end{array} = \Sigma\left[(x-\mu)(x-\mu)\right]$$

cor matrix

$$\rho = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_p \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \quad -1 \le \rho_{ij} \le 1$$

# Sampling



population
$X = [x_1, \ldots, x_p]^T$
$\mu$
$\Sigma$
random sample
$x_1, x_2, x_n$
$x_j \in \mathbb{R}^p$
estimate moments

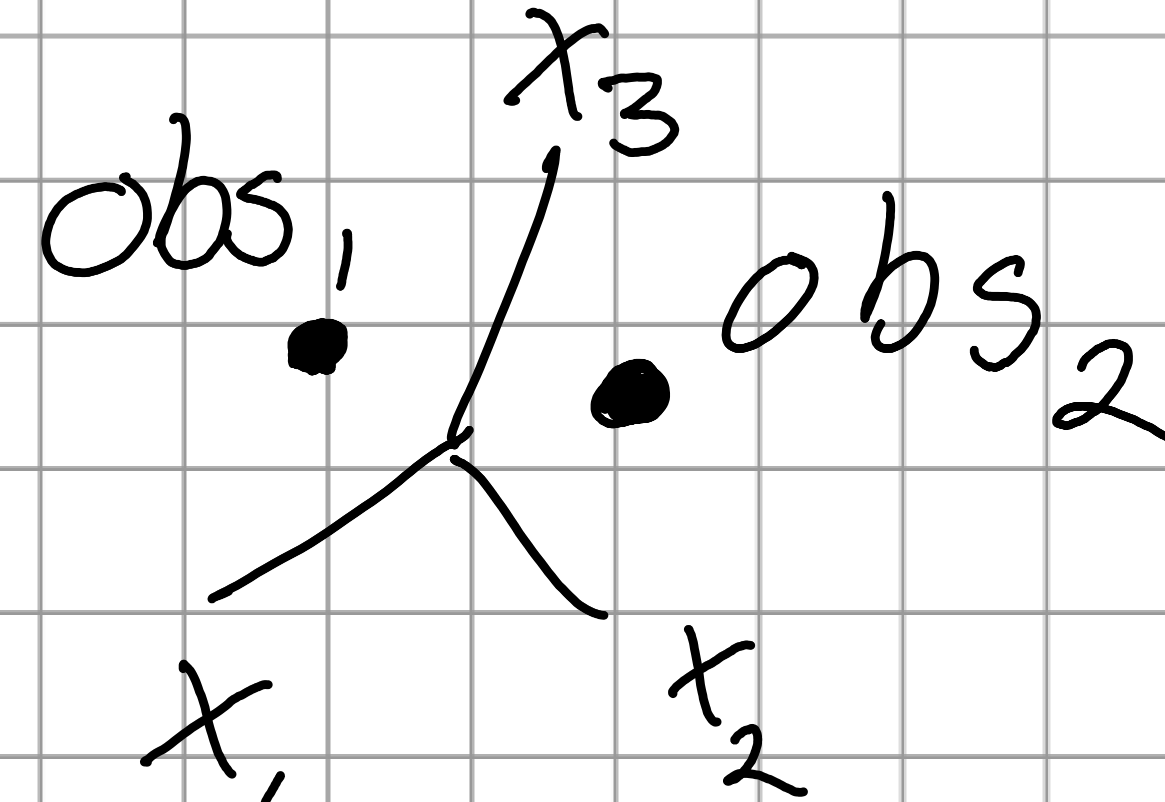
# Data Matrix

matrix

$$\underset{n\times p}{\underline{\underline{X}}} = \begin{array}{c} j=1 \\ \\ \\ \\ j=n \end{array} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & & & \\ \vdots & & & \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

$$= \begin{bmatrix} x_1^T \\ x_2^T \\ x_n^T \end{bmatrix}$$

n points in p-space



# Descriptiv Statistics

Estimation of moments $\begin{cases} \mu \\ \Sigma \\ \rho \end{cases}$

not Known, so we estimate

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

matlab: mean ($\bar{x}$)

Sum of all Samples

$$\underset{p\times p}{\hat{\Sigma}} = \underset{p\times p}{S} = \sum_{j=1} (x_j - \hat{\mu})(x_j - \hat{\mu})^T$$

$$= \begin{bmatrix} S_1^2 & S_{12} & S_{13} & \cdots & S_{1p} \\ & S_2^2 & & & \\ & & \ddots & & \\ -\,\|\,- & & & & S_p^2 \end{bmatrix}$$

$\bar{x}$, $S$, $R$

$$\hat{\rho} = R = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1p} \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ -\|- & & & & 1 \end{bmatrix}$$

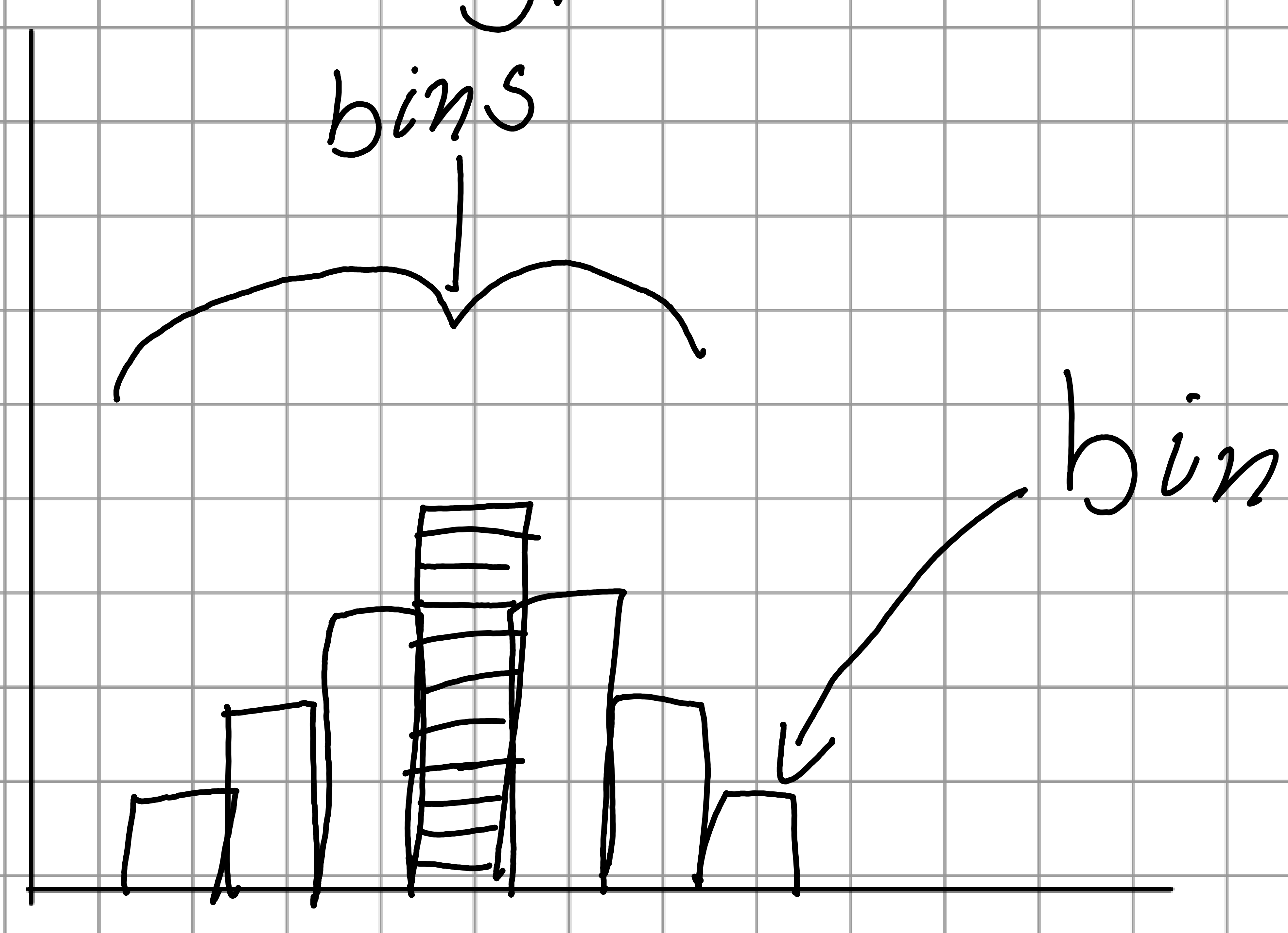
$$- \le r_{ij} \le 1$$

matlab: corrcoef($\bar{\bar{X}}$)

# Visualization of data

frequency / number of ocations
%

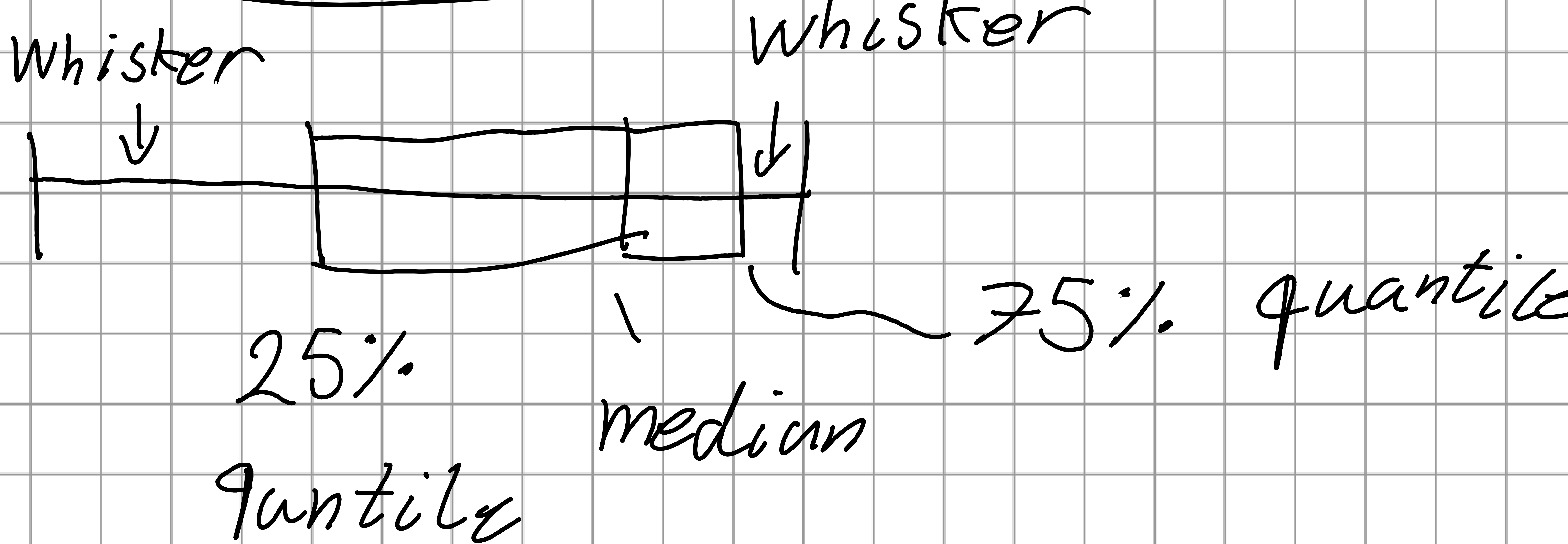
1D random variable, X
histogram



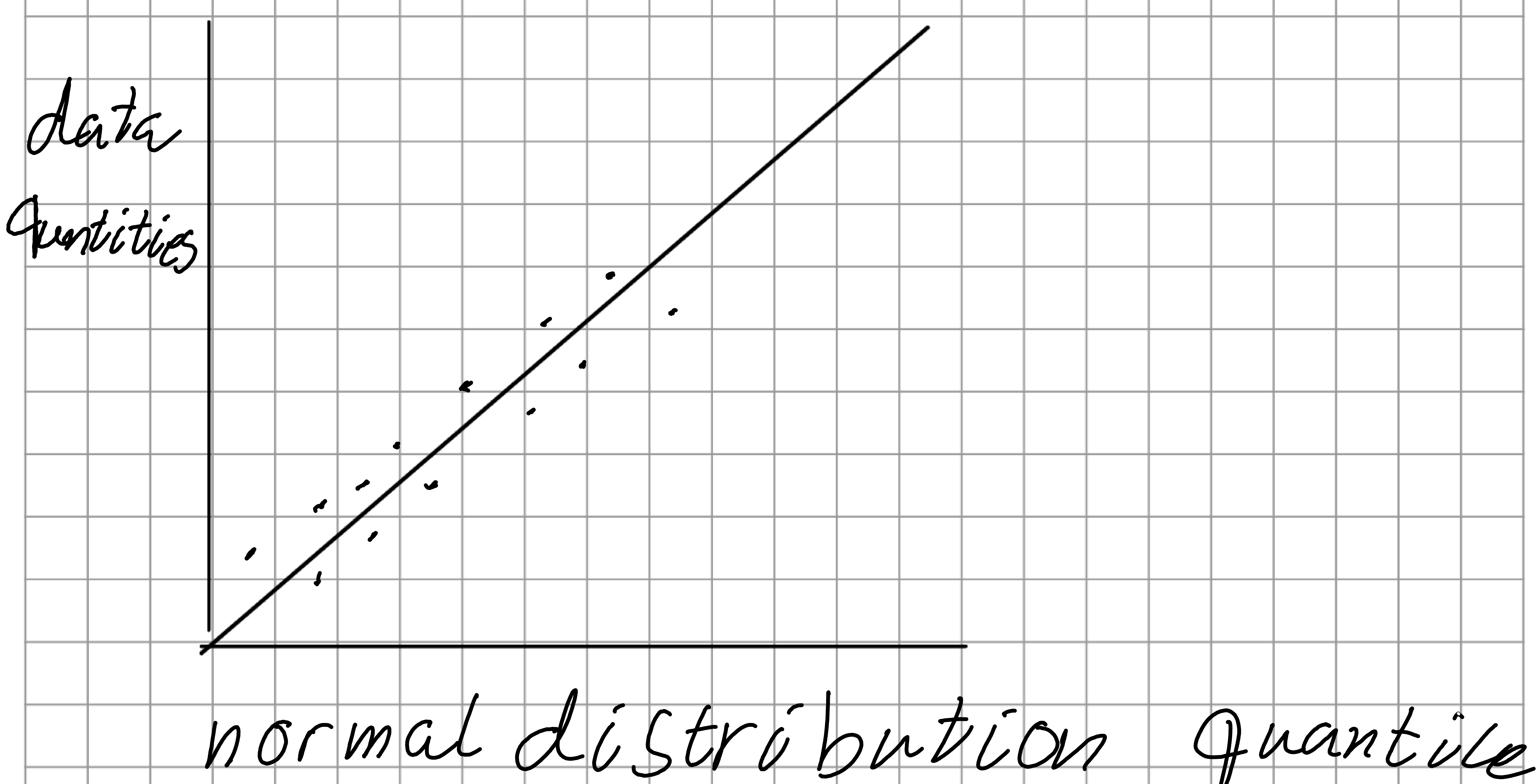

binning $\sqrt{n}$ = # bins

each "box" is an observation

# Box plot

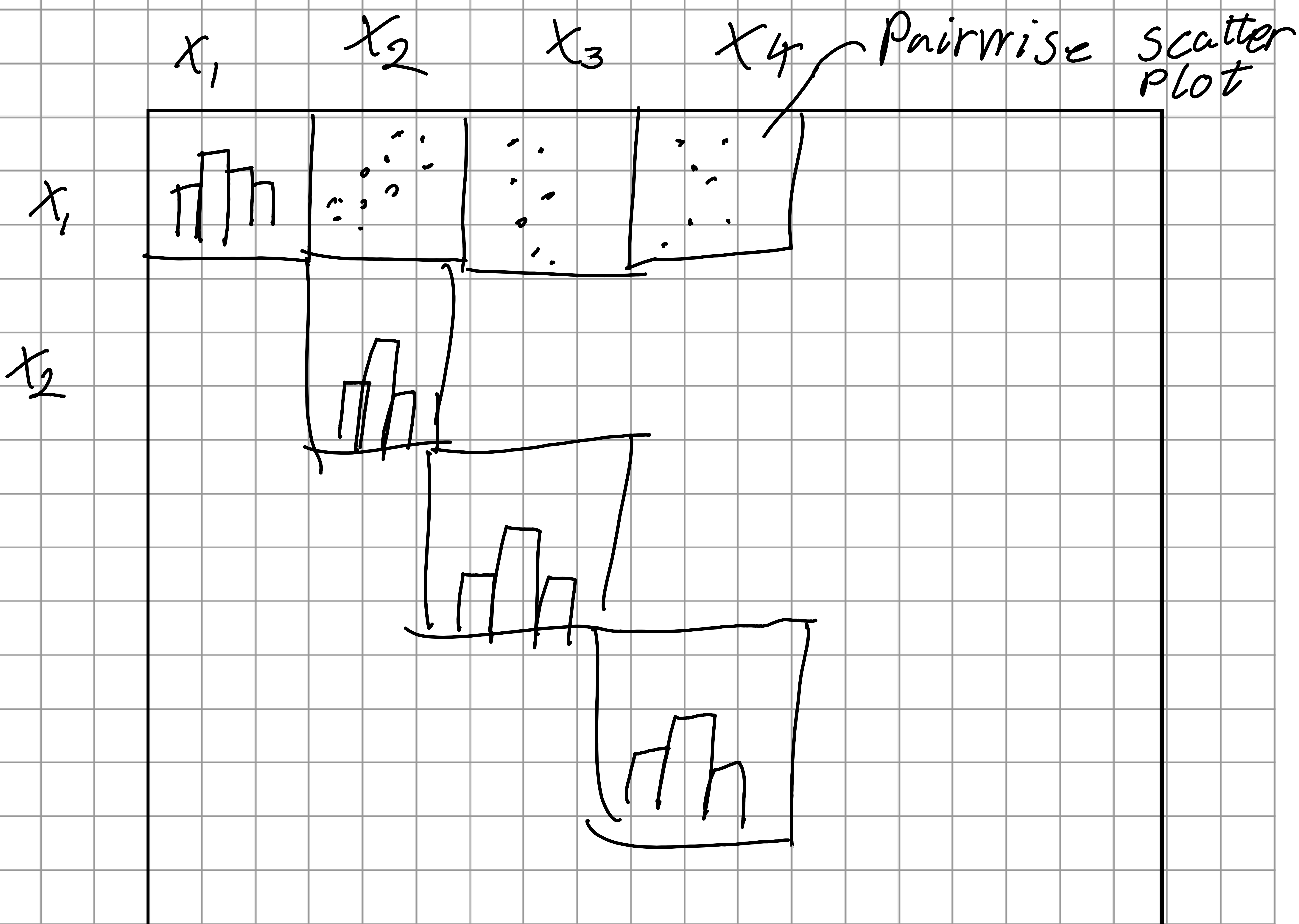



# QQ-plot



# 2D/pD matlab: potmatrix(X)

Scatter matrix




Only look at the diagonal