10/25/22

Scatterplot & vector plot

	X	Y
i	x_i	y_i

centralize $\longrightarrow$

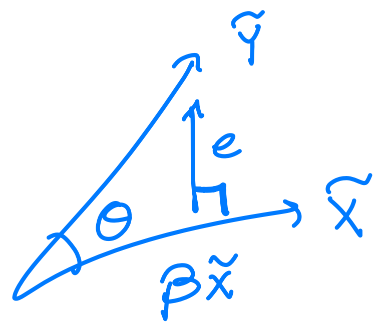
	$\tilde{X}$	$\tilde{Y}$
i	$x_i - \bar{x}$	$y_i - \bar{y}$

Scatterplot

row

2D

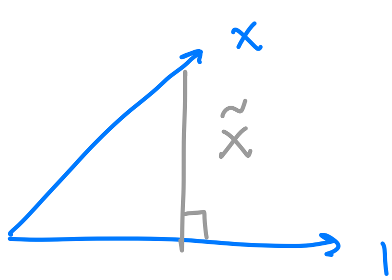
vector plot



n-dim

$$\rho = \mathrm{corr}(X, Y) = \cos\theta$$

$$1 - \rho^2 = \sin^2\theta = \frac{|e|^2}{|\tilde{y}|^2} \leftarrow \mathrm{var}\ e_i \quad = \frac{\mathrm{var}(e_i)}{\mathrm{var}(Y)}$$



	x	1
i	x_i	1

$$\vec{u} = \frac{1}{\sqrt{n}} 1$$

projection: $\langle x, \vec{u} \rangle \vec{u} = \frac{1}{n} \langle x, 1 \rangle 1 = \bar{x}\, 1$

$$\mathrm{var}(X) = E(X^2) - E(X)^2$$

$$\frac{1}{n}|x|^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \longrightarrow E(X^2)$$

$$\frac{1}{n}|\tilde{x}|^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 \longrightarrow \mathrm{Var}(X)$$

$$\frac{1}{n}|\bar{x}\mathbf{1}|^2 = \bar{x}^2 \longrightarrow E(X^2)$$

relate to pythagorean theorem

$$Z = X + Y$$

$$E(Z) = \int (x+y) f(x,y)\,dx\,dy$$

$$= E(X) + E(Y)$$

$$\mathrm{Var}(Z) = E((Z - E(Z))^2)$$

$$= E(((X+Y) - E(X+Y))^2)$$

$$= E(((X - E(X)) + (Y - E(Y)))^2)$$

$$= E((X - E(X))^2 + (Y - E(Y))^2 + 2(X - E(X))(Y - E(Y))$$

$$= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

$$X \perp Y \longrightarrow \mathrm{Cov}(X, Y) = 0$$

$$\mathrm{var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

$$X_1 \quad X_2 \; \ldots \; X_n \overset{iid}{\sim} f(x)$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

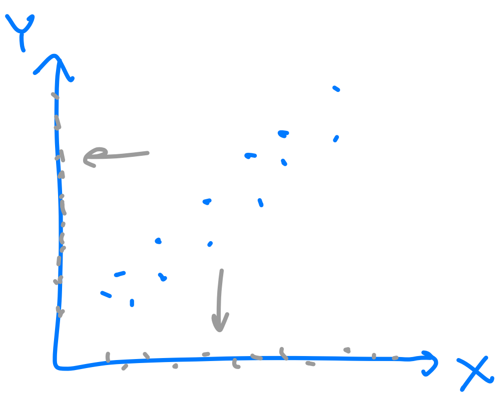
$$E(X_i) = \mu$$

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

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$$

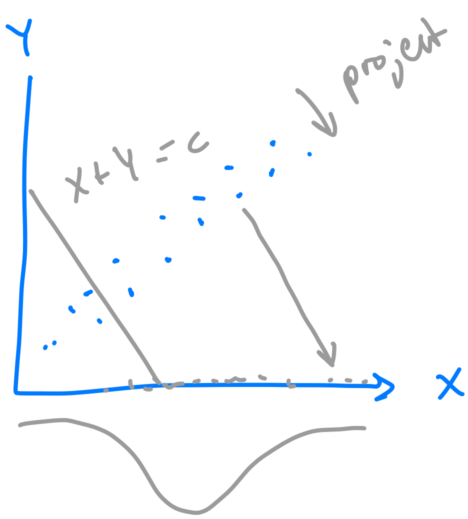
$$Var(\bar{X}) = \frac{1}{n^2}\underline{Var\left(\sum_{i=1}^{n} X_i\right)} = \frac{1}{n^2}\underline{\sum_{i=1}^{n} Var(X_i)} = \frac{\sigma^2}{n} \rightarrow$$

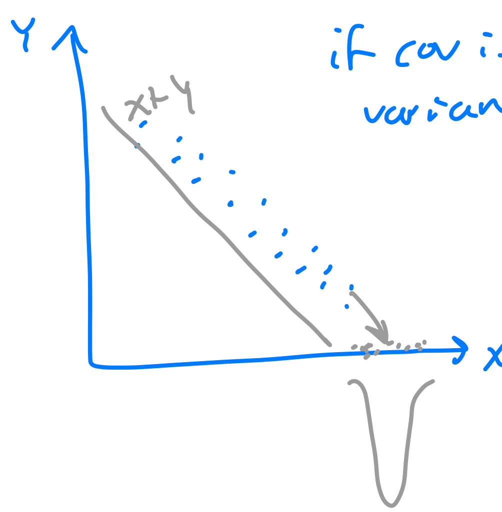
because of independence



marginal dist of y
↳ project to y axis

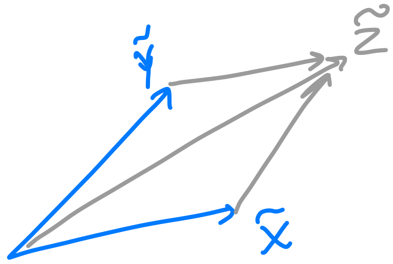
marginal dist of x
↳ project down

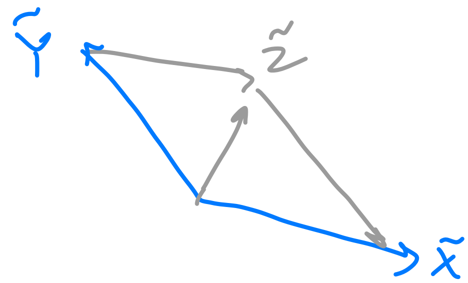




if cov is negative,
variance is small

variance of sums is controlled by covariance

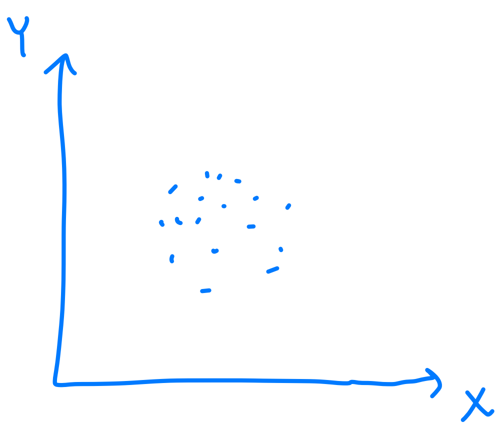


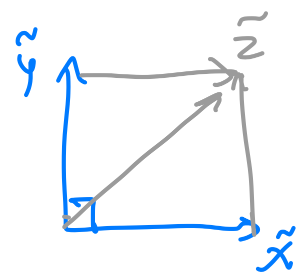


$$|\tilde{z}|^2 = |\tilde{x}|^2 + |\tilde{y}|^2 + 2|\tilde{x}||\tilde{y}|\cos\theta$$

$$\mathrm{var}(Z) = \mathrm{var}(X) + \mathrm{Var}(Y) + 2\mathrm{cov}(X, Y)$$

$$\langle \tilde{x}, \tilde{y} \rangle$$





$$|\tilde{z}|^2 = |\tilde{x}|^2 + |\tilde{y}|^2$$

$$\mathrm{var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Multivariate Statistics

Recall $X_{m \times n}$

$$\boxed{\begin{aligned} E(X) &= \left(E(X_{ij})\right)_{m \times n} \\ E(AX) &= A\,E(X) \\ E(XB) &= E(X) \cdot B \end{aligned}}$$

$$Var(X) = E\left((X-\mu)(X-\mu)^T\right)$$

$$\mu = E(X)$$

$$Var(AX) = E\left((AX - E(AX))(AX - E(AX))^T\right)$$

$$= A \, Var(X) \, A^T$$

$$Cov(X, Y) = E\left((X-\mu_X)(Y-\mu_Y)^T\right)$$

$d(x) \times d(y)$

$$Cov(AX, BY) = A \, Cov(X, Y) \, B^T$$

$$E(Y) = 0 \qquad Var(Y) = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_d \end{bmatrix}$$

$$E(y_i) = 0$$

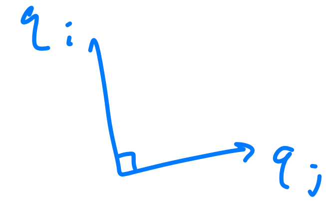
$$Var(Y) = \lambda_i \qquad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d > 0$$

$$Cov(y_i, y_j) = 0 \text{ for } i \neq j$$

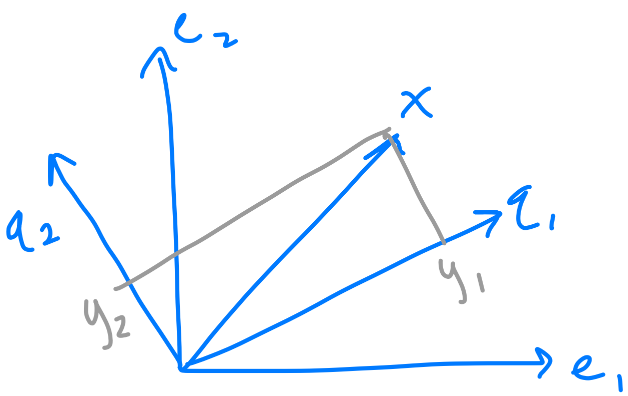
$$X_{d\times 1} = Q_{d\times d} Y_{d\times 1} = \begin{bmatrix} \vdots & & & & \vdots \\ q_1 & q_2 \cdots & q_i & \cdots & q_d \\ \vdots & & & & \vdots \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_d \end{bmatrix}$$

$$\langle q_i, q_j \rangle = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

orthonormal basis



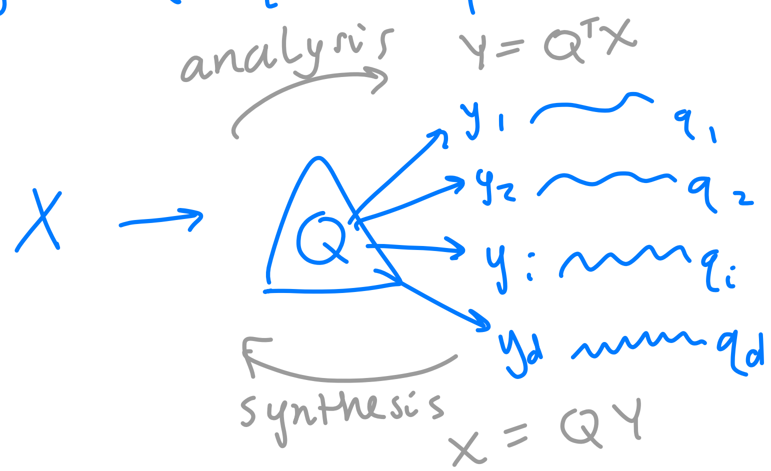
$$= y_1 q_1 + y_2 q_2 + \cdots + y_i q_i + \cdots y_d q_d = \sum_{i=1}^{d} y_i q_i$$



In math language: X becomes Y in basis Q

$$y_i = \langle x, q_i \rangle = q_i^T x \qquad \begin{bmatrix} \vdots \\ y_i \end{bmatrix} = \begin{bmatrix} q_i^T \end{bmatrix} x = Q^T X$$

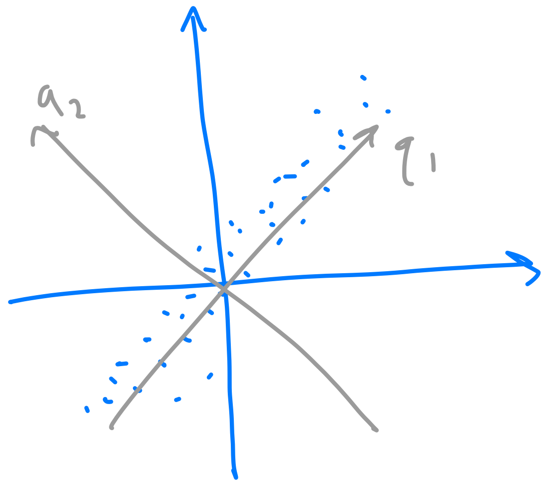
pretend number



$$E(X) = Q\,E(Y) = 0$$

$$\Sigma = Var(X) = Q\,Var(Y)\,Q^T = Q \Lambda Q^T$$

Spectral Theorem: $\Sigma = Q \Lambda Q^T$ → any symmetric matrix is a diagonal matrix

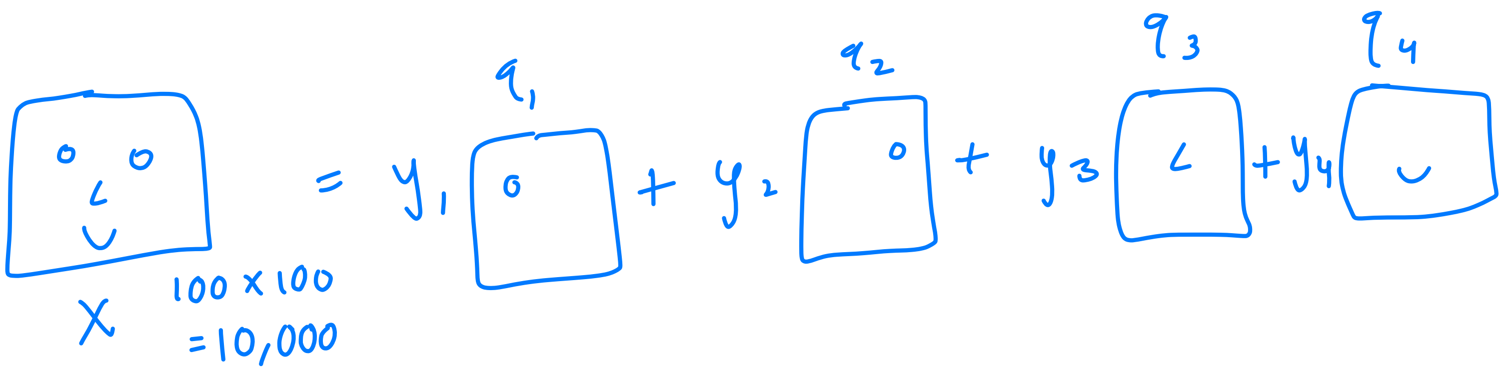


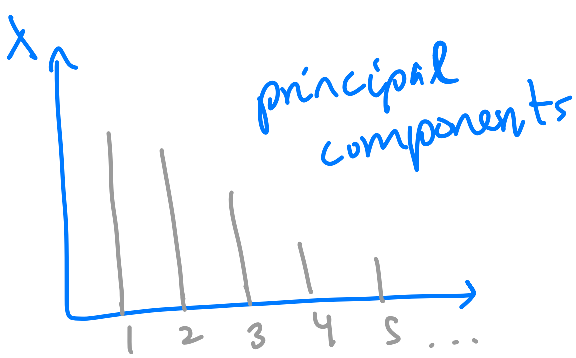
q_1 & q_2 are the new axes

↳ the points are uncorrelated

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \xrightarrow{Q} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

reduce to 1 dimension





after diagonalization, perform principal components analysis for dimension reduction

$$Z \sim N(0, I)$$

$E(z_i) = 0 \quad Var(z_i) = 1 \quad Cov(z_i, z_j) = 0 \quad i \neq j$

$$z_i \overset{iid}{\sim} N(0, 1)$$

$$f_z(z) = f(z_1, z_2 \dots z_i \dots z_d) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_i^2}{2}}$$

$$= \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{|z|^2}{2}}$$

Recall univariate

$$z = \frac{x-\mu}{\sigma} \qquad x = \mu + \sigma z$$

multivariate

$$X_{d\times 1} = \mu_{d\times 1} + \Sigma^{\frac{1}{2}}_{d\times d} Z_{d\times 1}$$

$$\Sigma^{\frac{1}{2}} = Q \Lambda^{\frac{1}{2}}$$ $\sqrt{\ }$ each diagonal element

$$= \mu + Q\Lambda^{\frac{1}{2}} Z$$

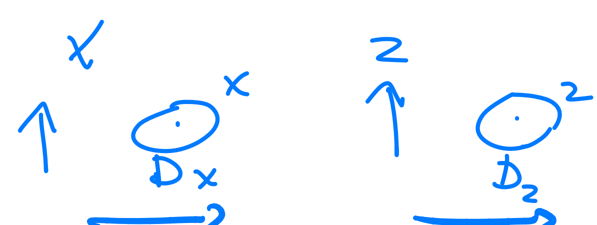
$$E(Z) = 0 \qquad var(Z) = I_d$$

$$E(X) = \mu \qquad var(X) = Q\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}Q^T = \Sigma$$

density of X

$$f_X(x)\,|D_x| = f_z(z)\,|D_z|$$

$$f_X(x) = f_z(z)\,\frac{1}{|D_x|/|D_z|} = \frac{|D_z|}{|D_x|}$$



det of Jacobian

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

$$\Sigma = Q\Lambda Q^T$$

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

change of variables

$$f_X(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{|z|^2}{2}} \quad \frac{1}{\det(\Sigma)^{\frac{1}{2}}}$$

$$= \frac{1}{(2\pi)^{\frac{d}{2}} \det(\Sigma)^{\frac{1}{2}}} e^{-\quad ?.}$$

$$z = \Lambda^{-\frac{1}{2}} Q^T (x-\mu)$$

$$|z|^2 = z^T z = (x-\mu)^T Q \Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{2}} Q^T (x-\mu)$$

$$= (x-\mu)^T \Sigma^{-1} (x-\mu)$$

$$X \sim N(\mu, \Sigma)$$

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

$$E(x) = \mu \quad Var(x) = \Sigma$$