MAESTRÍA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA

Universidad Tecnológica de Pereira

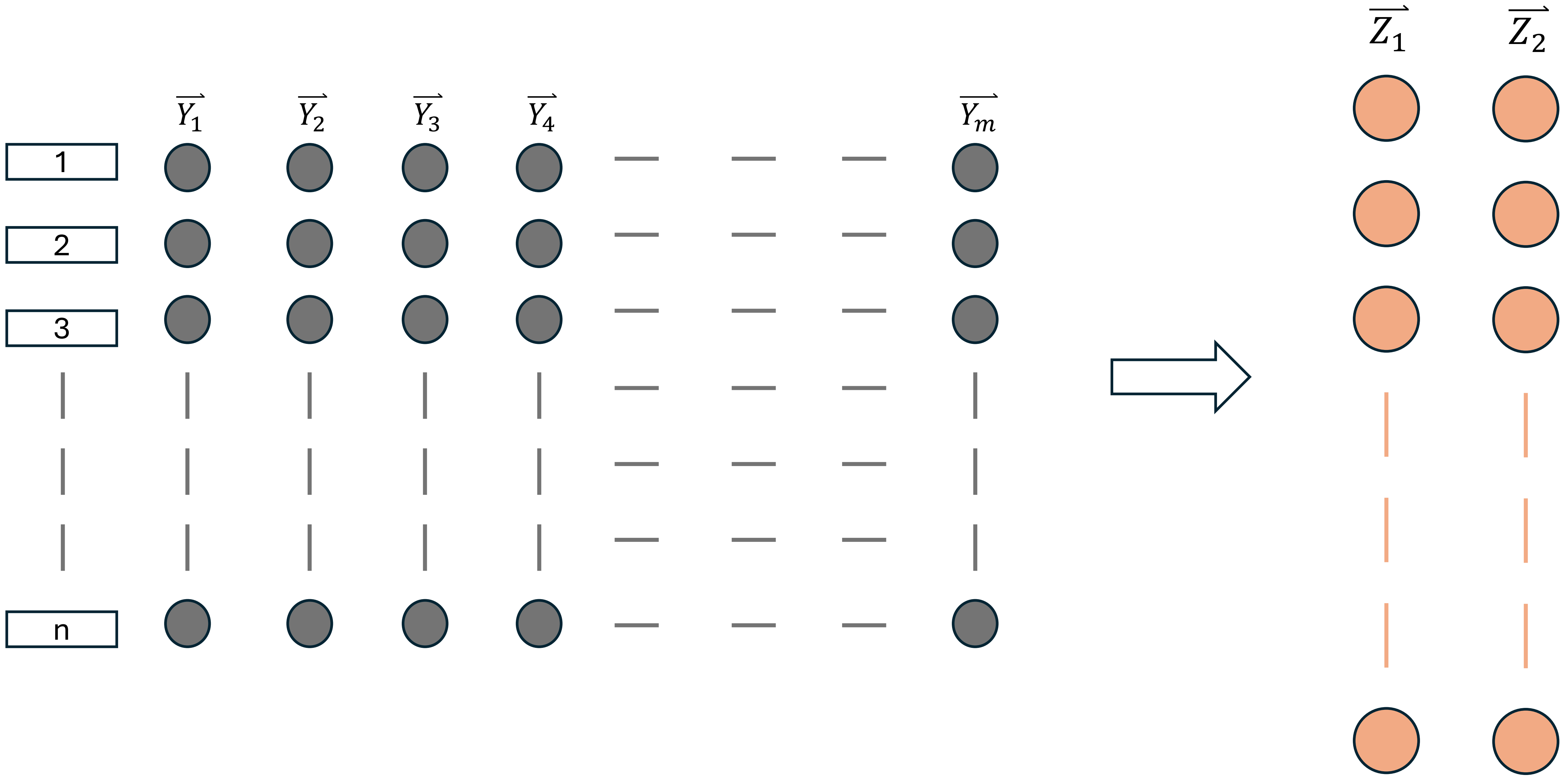
Multivariate statistics

Principal Component Analysis
PCA

May 18, 2024

Orlando Sabogal-Cardona
@Antonio Sabogal
orlando.sabogal.20@ucl.ac.uk

General idea



Do you know what a variate is?
A linear combination?

The variate:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_p X_p$$

$$Y_{mi} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_{pi} X_{pi}$$

n: $number\ of\ observations\ (individuals)$

p: $number\ of\ original\ variables$

m: $number\ of$ "new" variables

m: projections, component, factors

The variate:

$$Y = X\beta$$

The variate:

$$Y = X\beta$$

$$Y = \begin{matrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_m \end{matrix} \qquad \beta = \begin{matrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{matrix} \qquad X = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ 1 & X_{13} & X_{23} & \cdots & X_{p3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix}$$

Eigenvalues and eigenvectors

An eigenvector "v" is non-zero vector that satisfies the following equation:

$$Av = \lambda v$$

If we multiply matrix "A" by vector "v", the new vector does not change the direction after the transformation

The eigenvalue tells how much the eigenvector changes when multiplied by a matrix. It is an scaling value

Given this matrix:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$$

Is this vector:

$$v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

An eigenvector of "A"?

Given this matrix: $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$

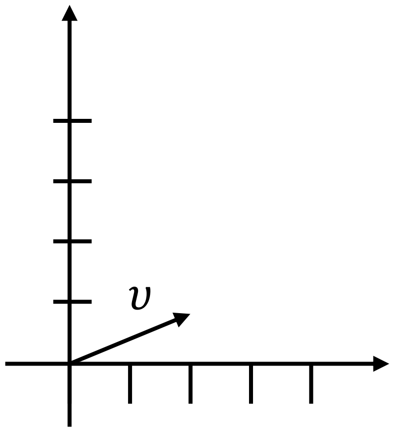
Is this vector: $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

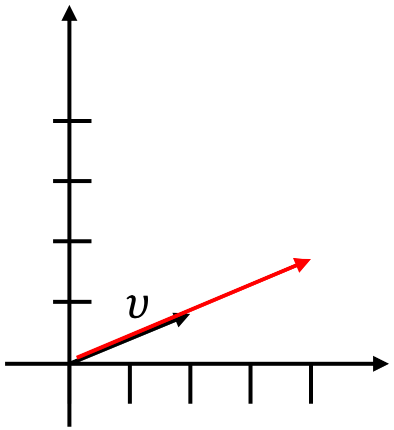
An eigenvector of "A"?

$$Av = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} * \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$$\lambda = 2$$

$$Av = 2 * \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$





A matrix can have multiple Eigenvectors.
Nevertheless, eigenvalue does not change.

Given this matrix: $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$

Is this vector: $v = \begin{bmatrix} 3 \\ 1.5 \end{bmatrix}$

An eigenvector of "A"?

What about this:

Given this matrix: $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$

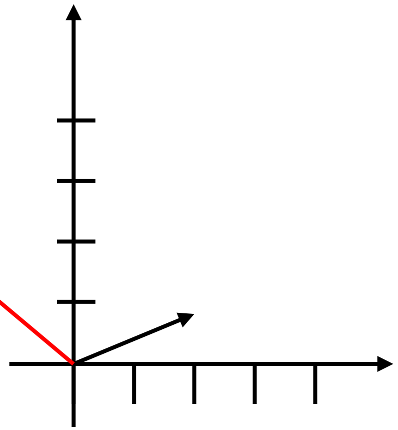
Is this vector: $v = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$

An eigenvector of "A"?

What about this:

Given this matrix: $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ Is this vector: $v = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$ An eigenvector of "A"?

$$Av = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} * \begin{bmatrix} -3 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix} \qquad \lambda = -1 \qquad Av = -1 * \begin{bmatrix} -3 \\ 3 \end{bmatrix}$$



How to calculate the "Eigens"?

$$Av = \lambda v \longrightarrow Av = \lambda I v$$

I: *identity matrix, ones in the diagonal*

$$Av - \lambda I v = 0 \longrightarrow \underbrace{(A - \lambda I)}_{M} v = 0$$

If there is a vector that, after multiplying a matrix, produces zero, the matrix is not invertible

The determinant of a non-invertible matrix is zero.

$$det(A - \lambda I) = 0$$

$$det(A - \lambda I) = 0 \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$det(A - \lambda I) = 0 \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0 \longrightarrow det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0$$

$$det(A - \lambda I) = 0 \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0 \longrightarrow det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0$$

$$det\left(\begin{bmatrix} 1 - \lambda & 2 \\ 1 & -\lambda \end{bmatrix}\right) = 0 \longrightarrow -\lambda + \lambda^2 - 2 = 0 \longrightarrow \lambda^2 - \lambda - 2 = 0$$

$$det(A - \lambda I) = 0 \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0 \longrightarrow det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0$$

$$det\left(\begin{bmatrix} 1-\lambda & 2 \\ 1 & -\lambda \end{bmatrix}\right) = 0 \longrightarrow -\lambda + \lambda^2 - 2 = 0 \longrightarrow \lambda^2 - \lambda - 2 = 0$$

$$\boxed{(\lambda - 2)(\lambda + 1) = 0}$$

$$det(A - \lambda I) = 0 \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0 \longrightarrow det\left(\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0$$

$$det\left(\begin{bmatrix} 1 - \lambda & 2 \\ 1 & -\lambda \end{bmatrix}\right) = 0 \longrightarrow -\lambda + \lambda^2 - 2 = 0 \longrightarrow \lambda^2 - \lambda - 2 = 0$$

$$\boxed{(\lambda - 2)(\lambda + 1) = 0}$$

$$\lambda = 2$$

$$\lambda = -1$$

How to get the Eigenvectors?

$$Av = \lambda v \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad \lambda = 2$$

$$\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 2 \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$Av = \lambda v \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad \lambda = 2$$

$$\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 2 \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$z_1 + 2z_2 = 2z_1$$

$$z_1 = 2z_2$$

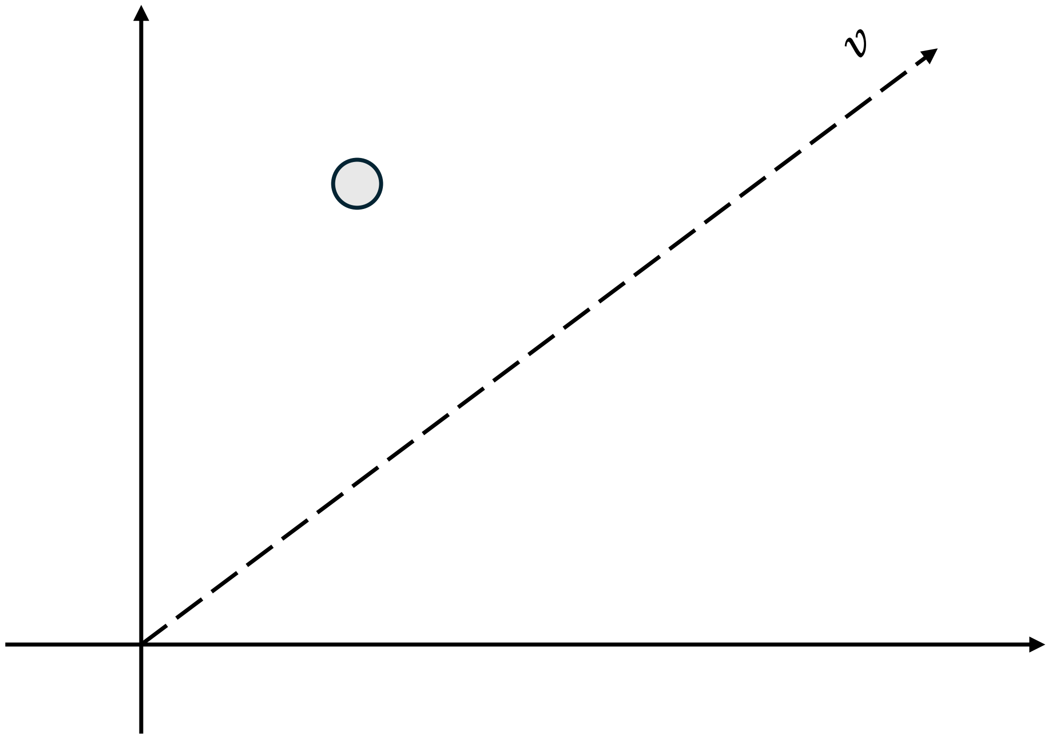
$$Av = \lambda v \qquad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \qquad \lambda = 2$$

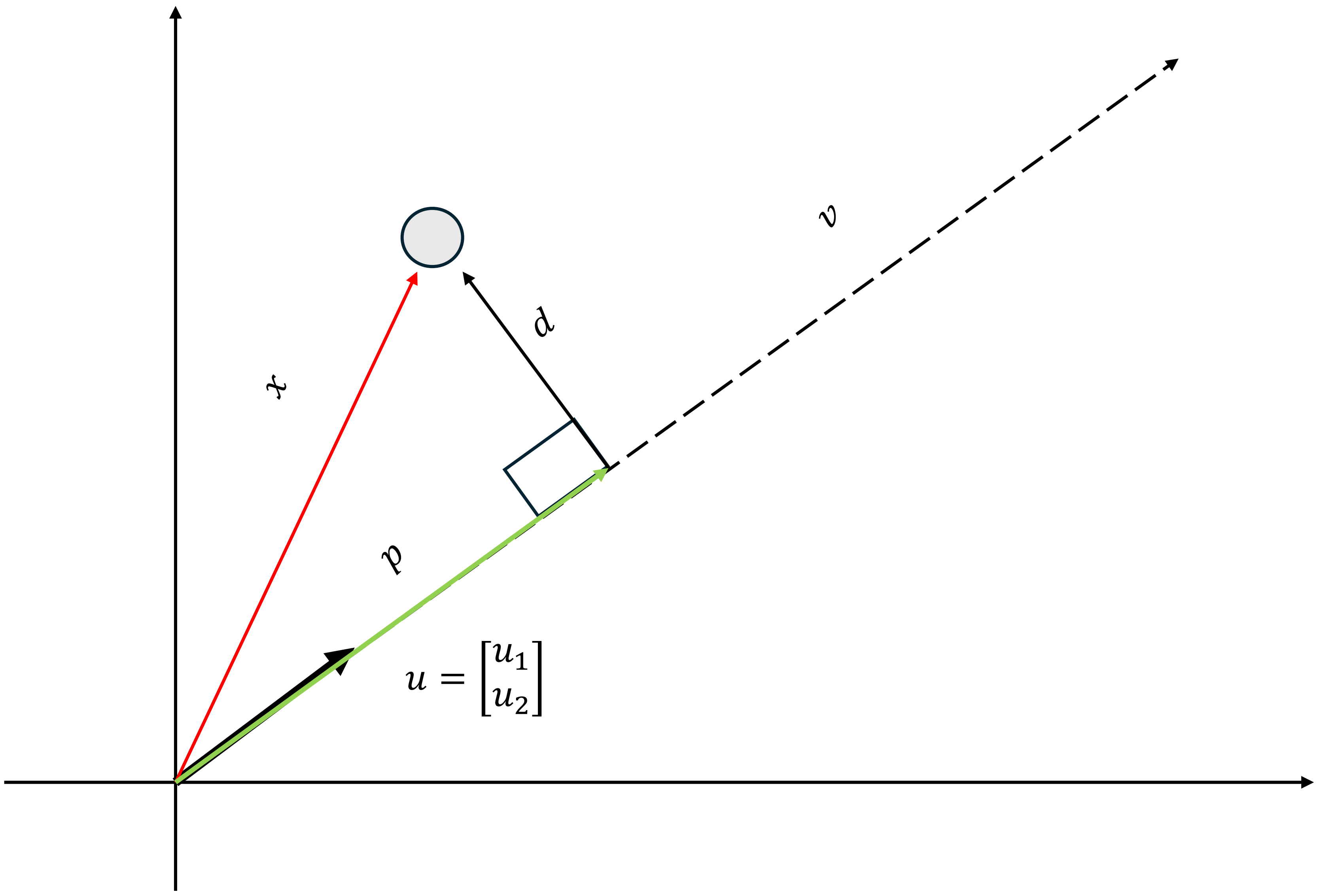
$$\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 2 \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$z_1 + 2z_2 = 2z_1 \longrightarrow 2z_2 = 2z_1 - z_1 \longrightarrow 2z_2 = z_1$$

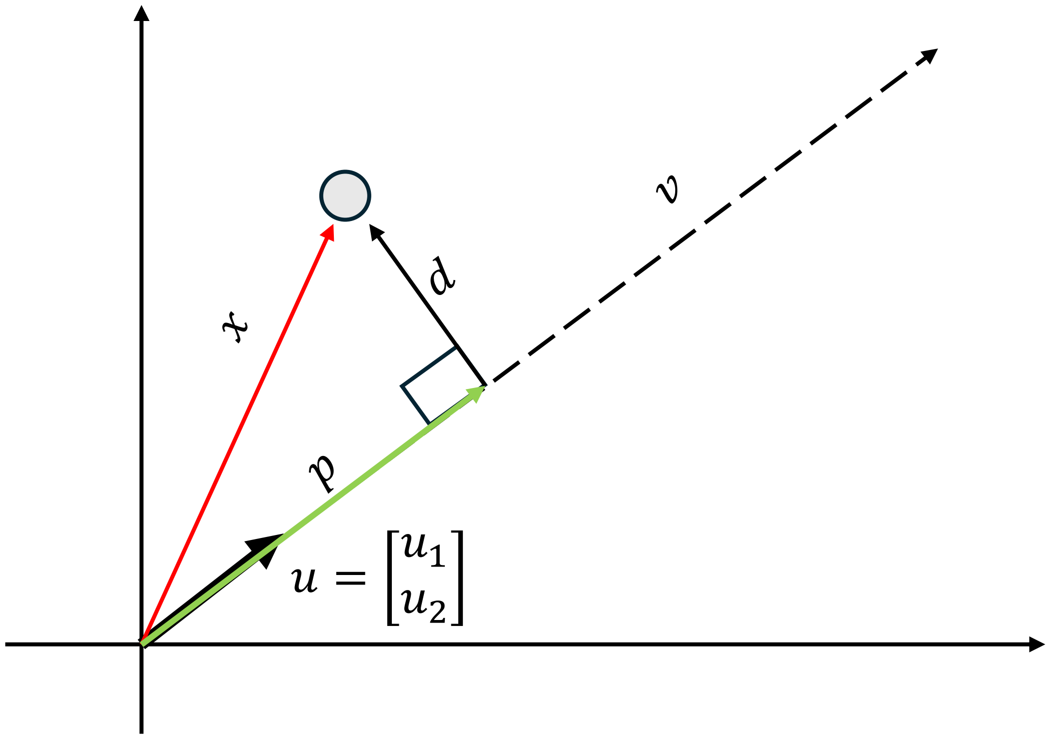
$$z_1 = 2z_2$$

Vector projections





x
d
v
p
$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$



u: *Unit Vector*

Dot product of a vector with itself is equal to its magnitude.

$$u = \frac{v}{||v||}$$

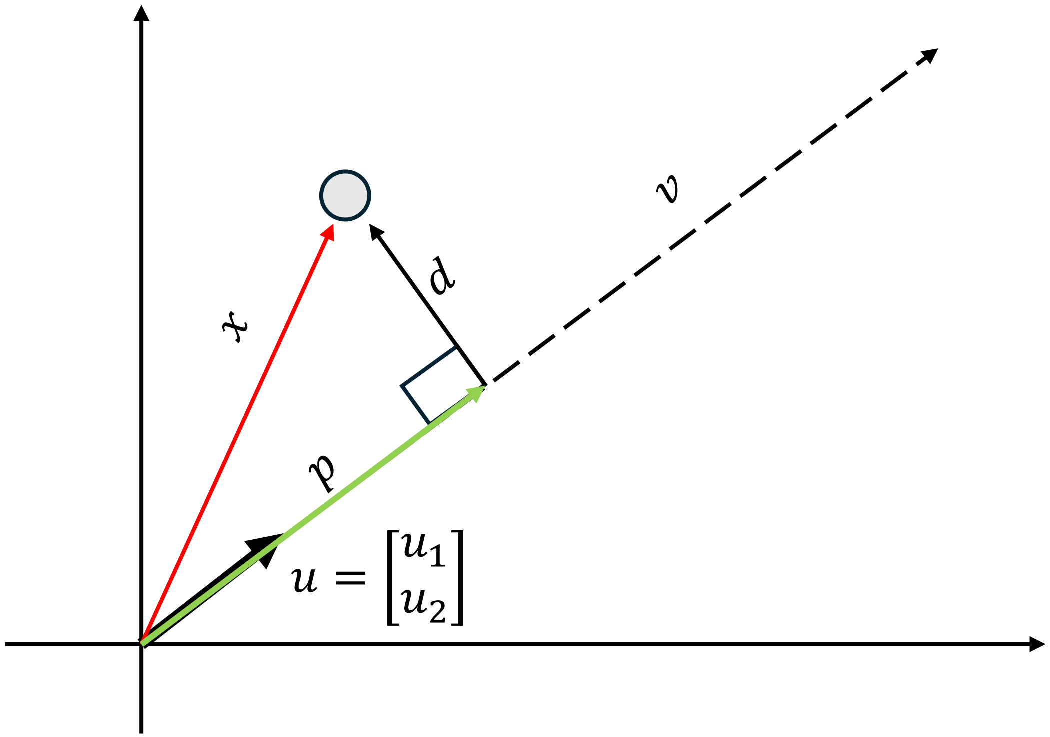
$$u \cdot u = 1$$

$$p = ku$$

$$p + d = x$$

$$d = x - p$$

$$d = x - ku$$



u: *Unit Vector*

Dot product of a vector with itself is equal to its magnitude.

$$u = \frac{v}{||v||} \qquad u \cdot u = 1$$

$$p = ku$$

$$p + d = x$$

$$d = x - p$$

$$d = x - ku$$

Remember: The dot product of two perpendicular (orthogonal) vectors is zero.

$$\longrightarrow p \cdot d = 0 \longrightarrow (ku) \cdot (x - ku) = 0 \longrightarrow kux - kuku = 0$$

$$\longrightarrow k(ux - kuu) = 0 \longrightarrow ux - kuu = 0 \longrightarrow ux - k = 0$$

$$\longrightarrow k = xu \longrightarrow p = (x \cdot u)u \longrightarrow p = \frac{x \cdot v}{||v||}u \longrightarrow \boxed{p = u^T x\, u}$$

PCA: long story short

$$SZ = \lambda Z$$

S: Covariance Matrix

Z: Eigenvectors

λ: *Eigenvalues*

Standardize original variables $\frac{X - \mu}{\sigma}$

↓

Compute "S"

↓

Calculate eigenvalues and eigenvectors

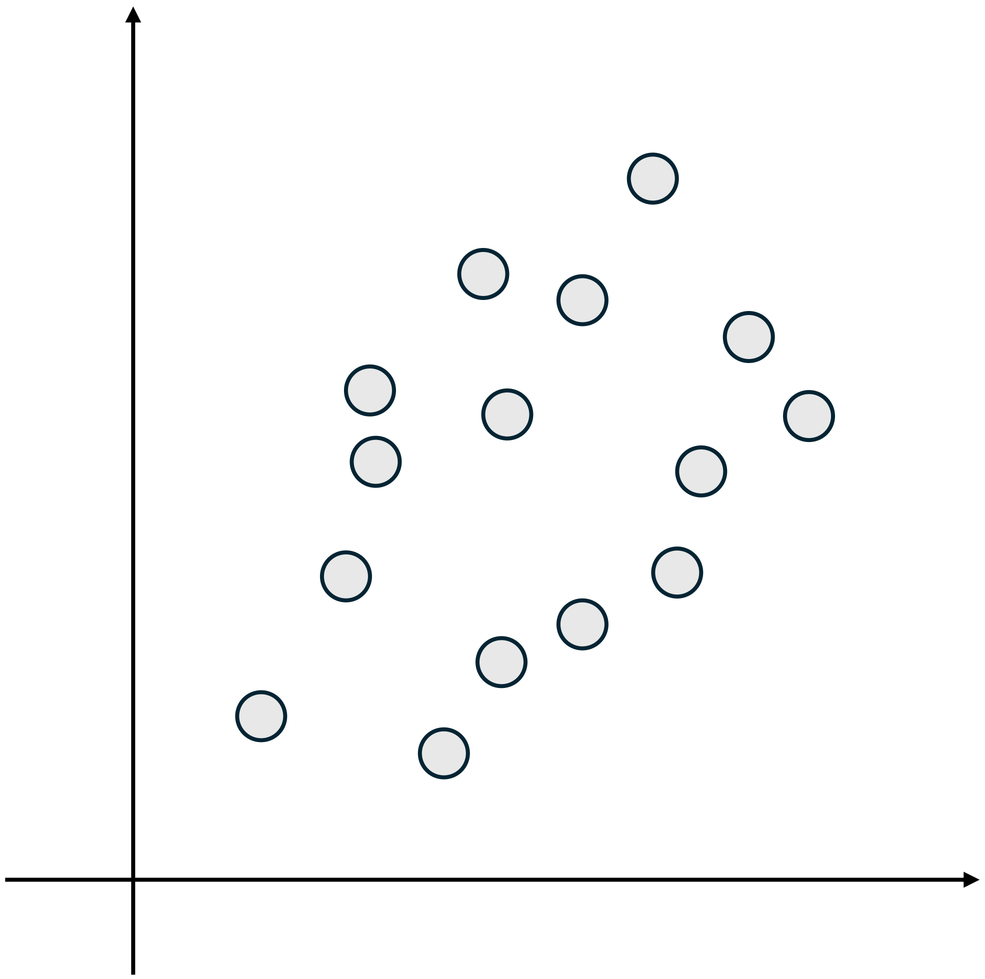
↓

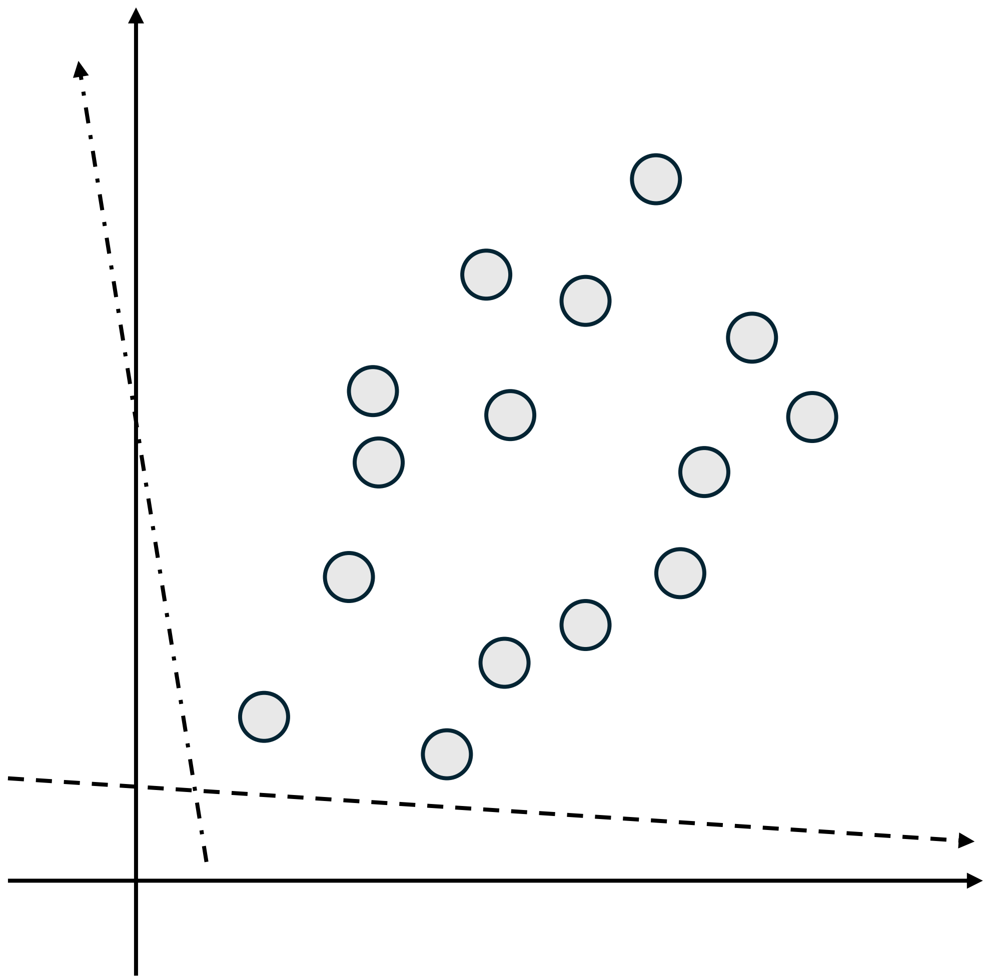
Sort by Eigenvalues: they explain variance

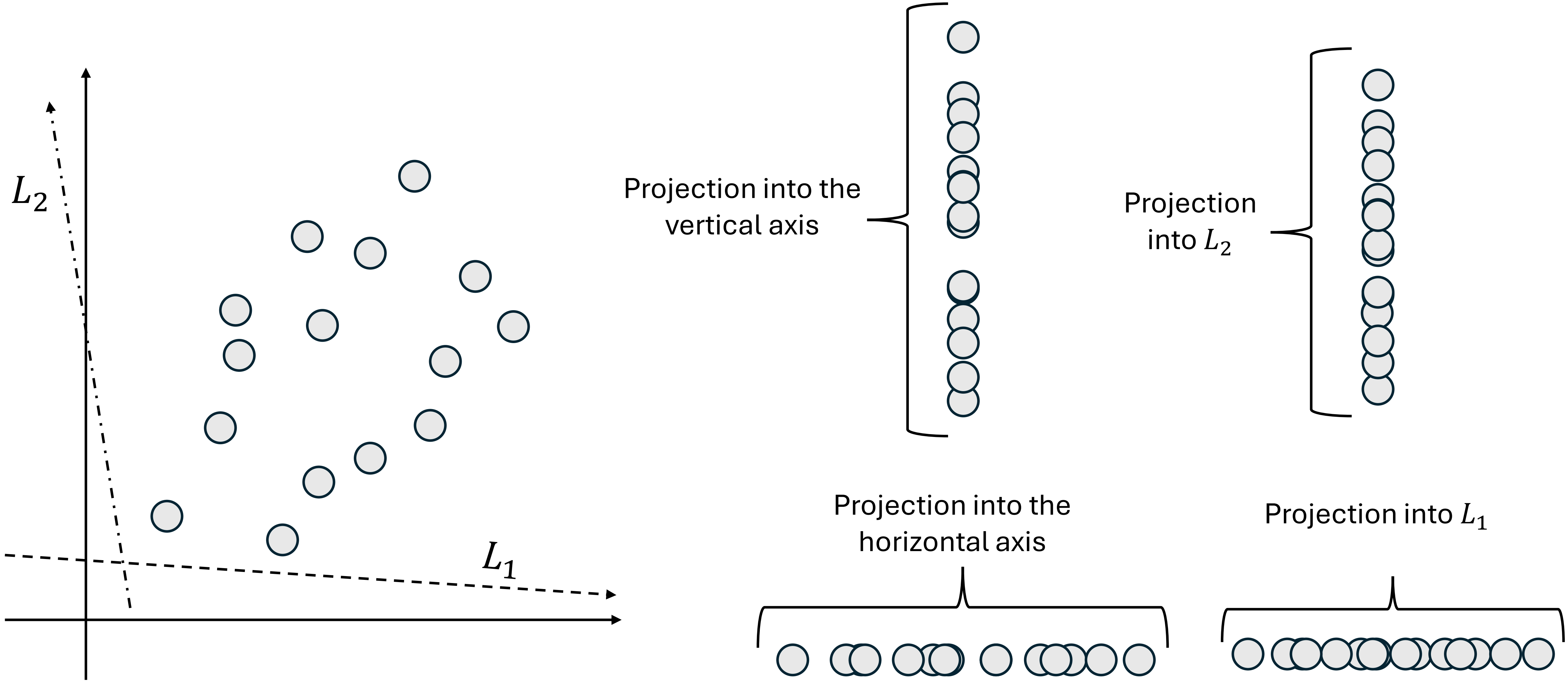
↓

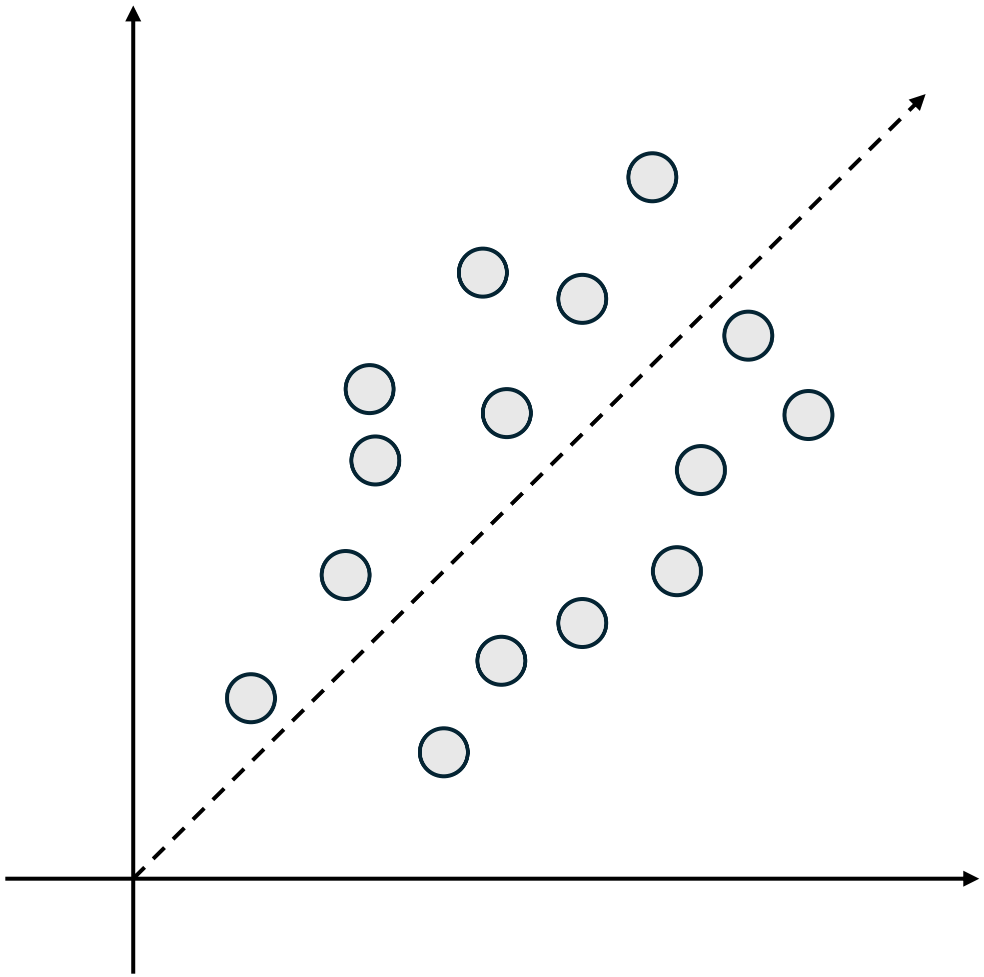
Transform (reproject) and interpret

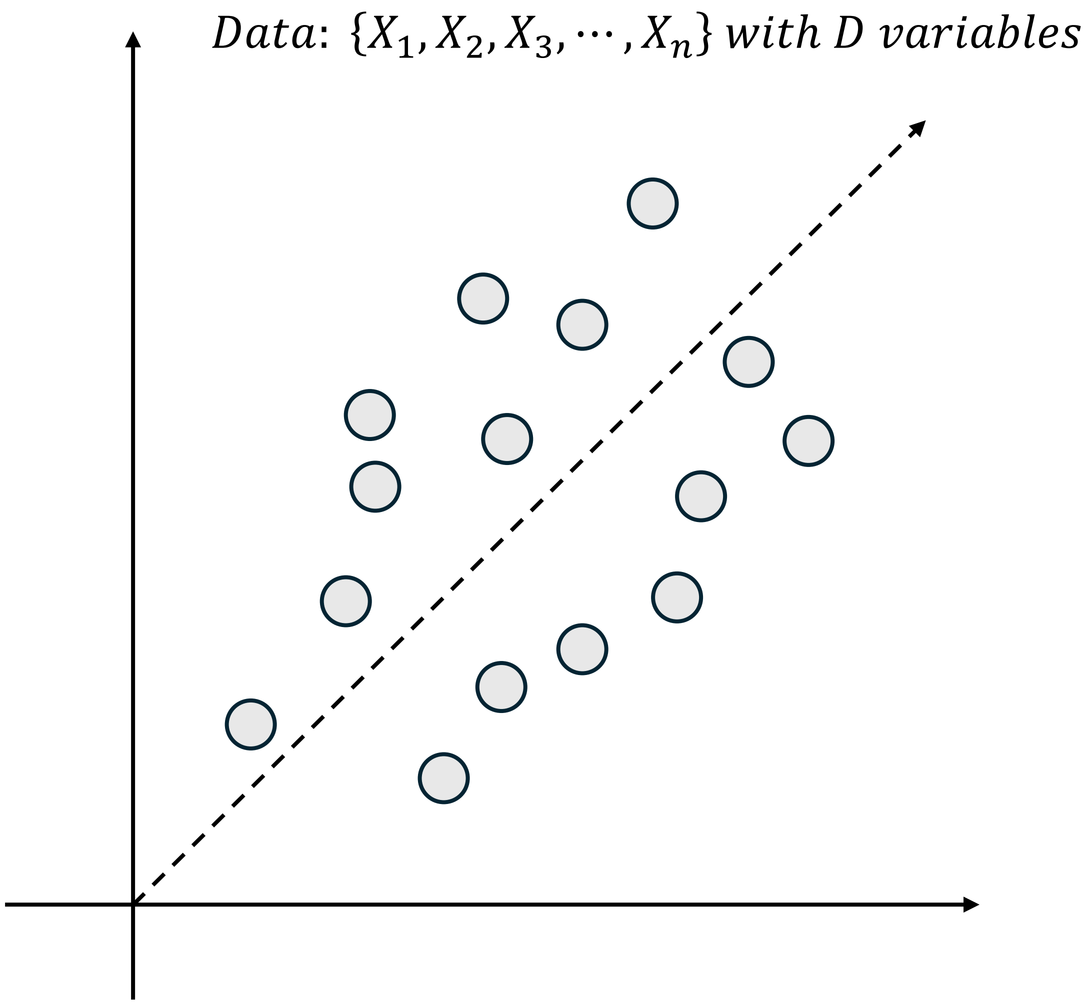
Where and how would you project this data?
(you want to preserve as much as the original variability)






L_2
L_1
Projection into the vertical axis
Projection into L_2
Projection into the horizontal axis
Projection into L_1



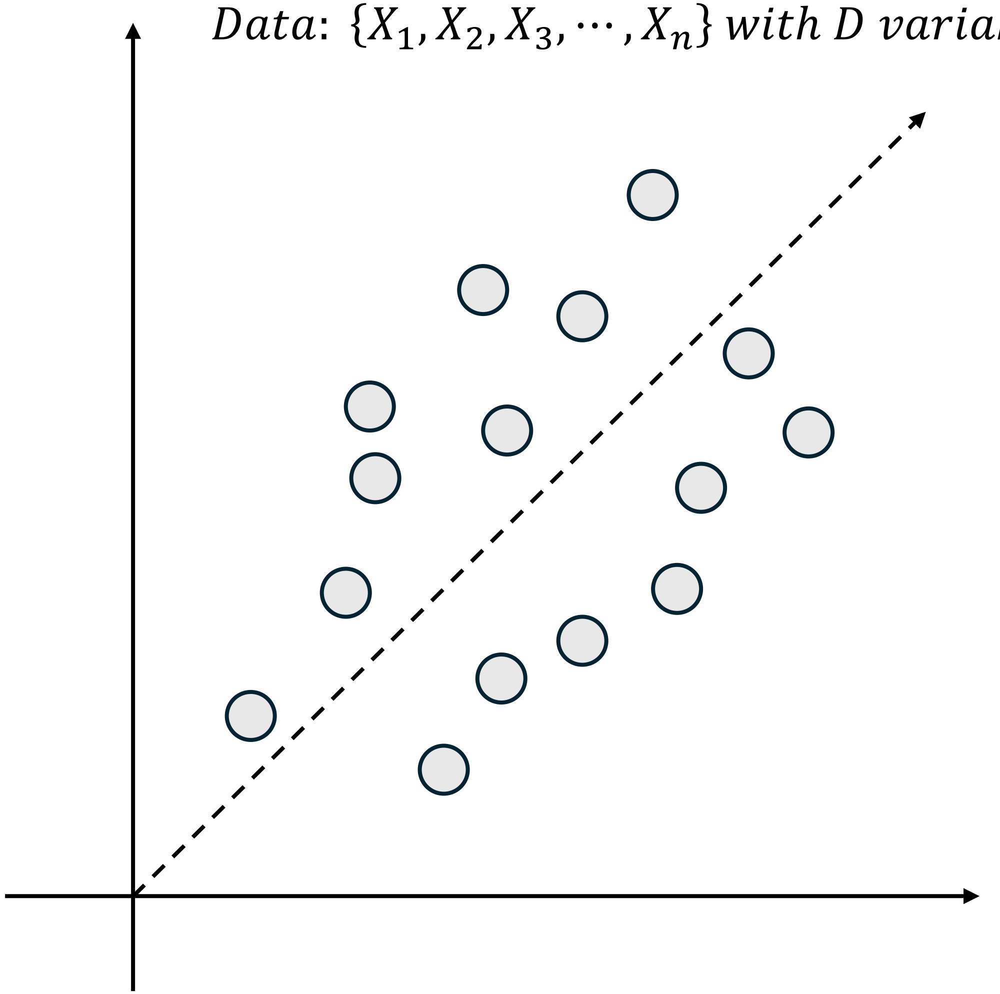


$$PROJE_{u_1}(x_i) = u_1^T x_i u$$

$$PROJE_{u_1}(\bar{x}) = u_1^T \bar{x}\, u$$

u: *Unit Vector*

How to get a variance of the projections?



Data: $\{X_1, X_2, X_3, \cdots, X_n\}$ with D variables

$$PROJE_{u_1}(x_i) = u_1^T x_i u$$

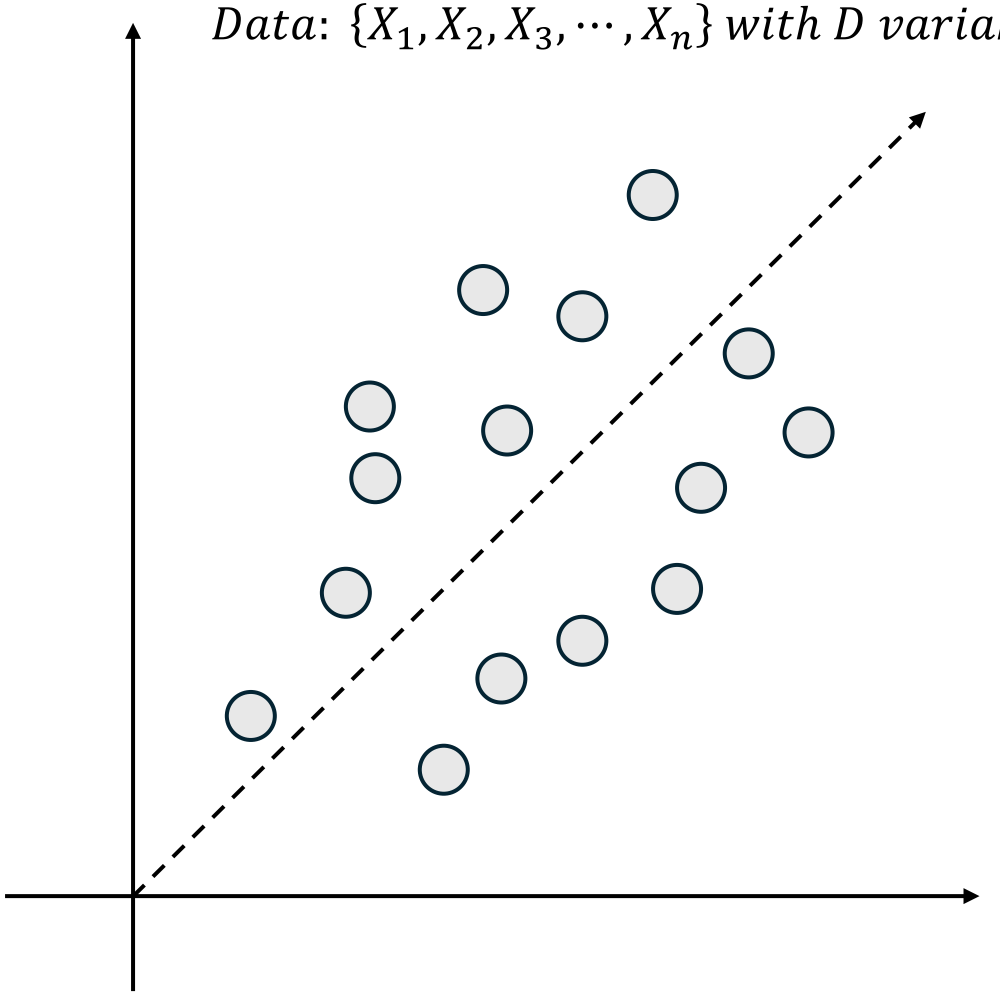
$$PROJE_{u_1}(\bar{x}) = u_1^T \bar{x}\, u$$

u: Unit Vector

How to get a variance of the projections?

$$\boxed{\frac{1}{n}\sum_{i=1}^{n}(u_1^T x_i - u_1^T \bar{x})^2}$$

$Data: \{X_1, X_2, X_3, \cdots, X_n\}$ $with\ D\ variables$



$$PROJE_{u_1}(x_i) = u_1^T x_i u$$

$$PROJE_{u_1}(\overline{x}) = u_1^T \overline{x}\, u$$

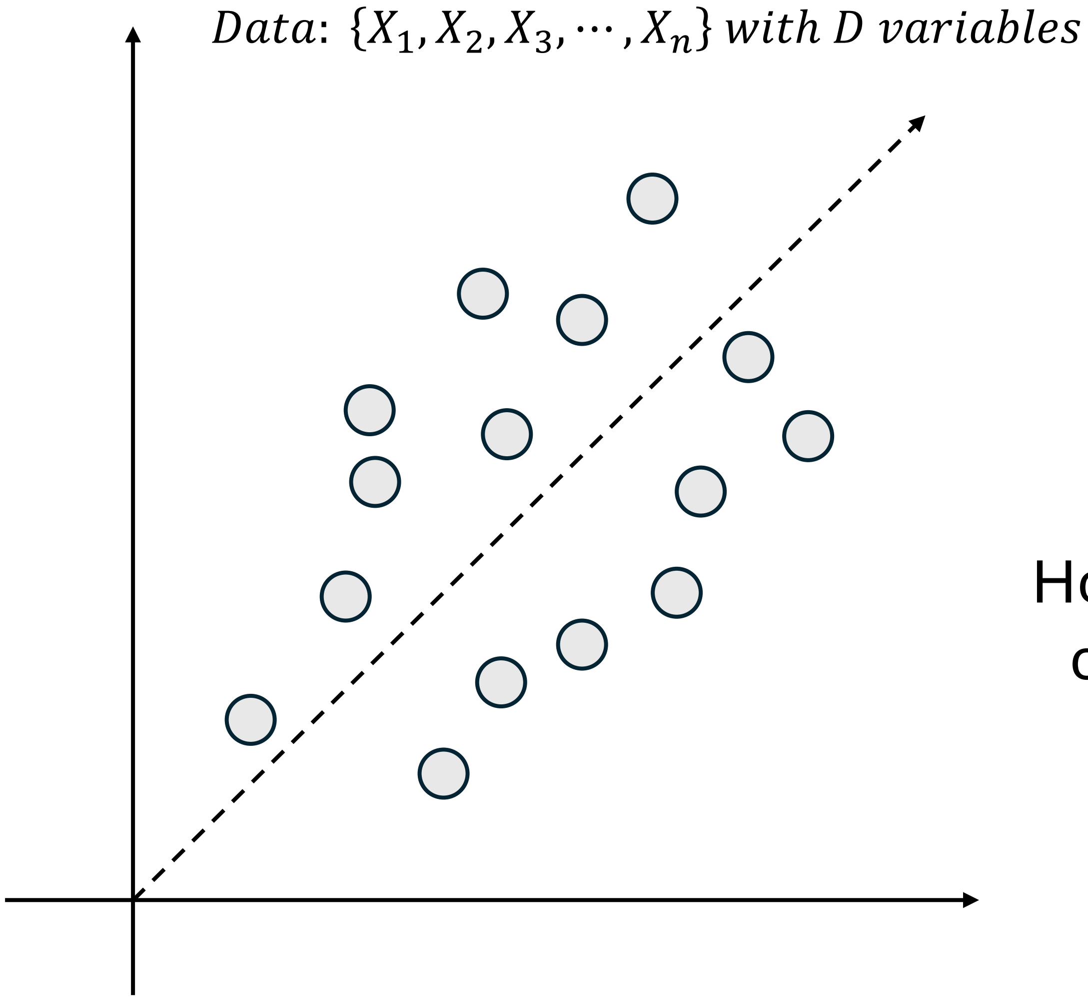
u: $Unit\ Vector$

How to get a variance of the projections?

$$\boxed{\frac{1}{n}\sum_{i=1}^{n}(u_1^T x_i - u_1^T \overline{x})^2}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(u_1^T (x_i - \overline{x})\right)^2$$

$$\frac{1}{n}\sum_{i=1}^{n} u_1^T (x_i - \overline{x})(x_i - \overline{x})^T u_1 \longrightarrow u_1^T \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(x_i - \overline{x})^T u_1$$



$$PROJE_{u_1}(x_i) = u_1^T x_i u$$

$$PROJE_{u_1}(\bar{x}) = u_1^T \bar{x}\, u$$

u: *Unit Vector*

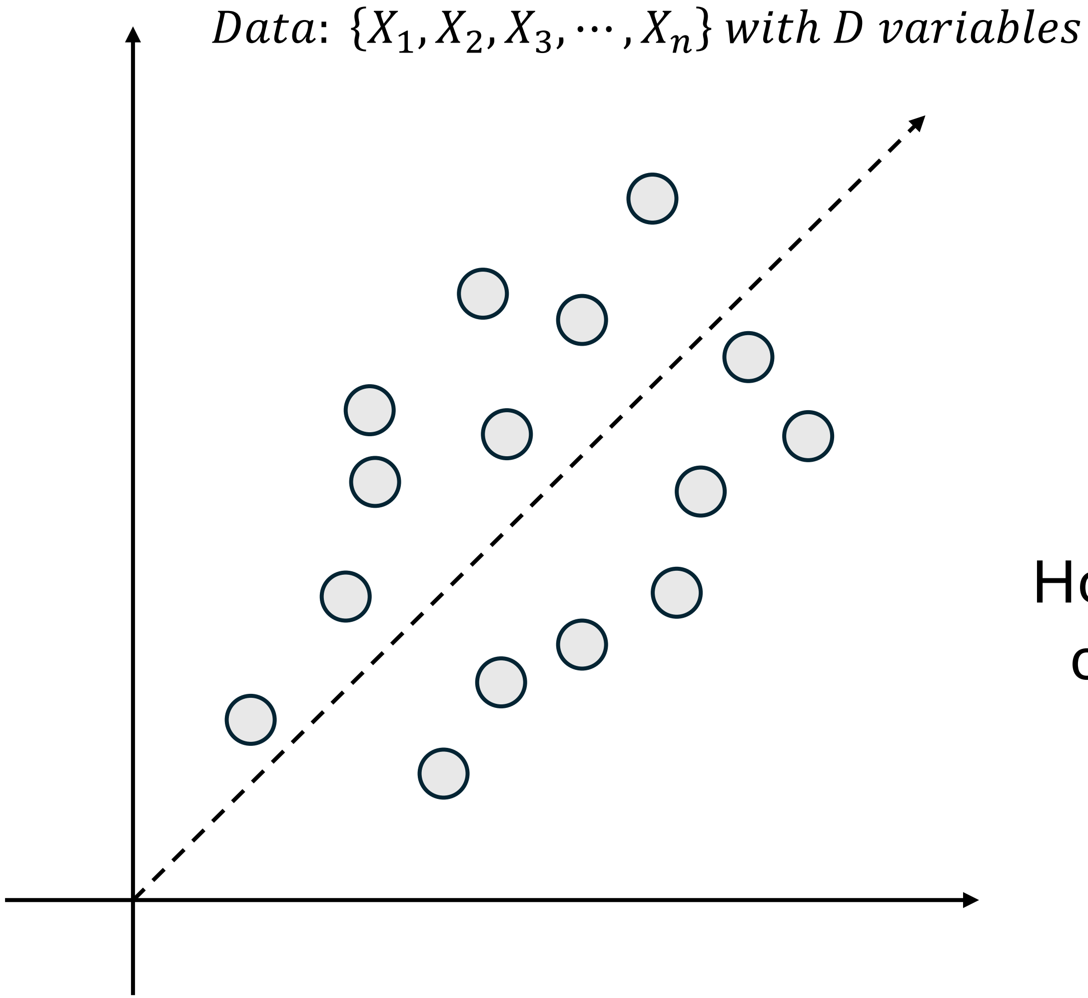
How to get a variance of the projections?

$$\boxed{\frac{1}{n}\sum_{i=1}^{n}(u_1^T x_i - u_1^T \bar{x})^2}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(u_1^T (x_i - \bar{x})\right)^2$$

$$\frac{1}{n}\sum_{i=1}^{n} u_1^T (x_i - \bar{x})(x_i - \bar{x})^T u_1 \longrightarrow u_1^T \boxed{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T} u_1$$

What is this? Does it ring a bell?



$$PROJE_{u_1}(x_i) = u_1^T x_i u$$

$$PROJE_{u_1}(\bar{x}) = u_1^T \bar{x}\, u$$

u: $Unit\ Vector$

How to get a variance of the projections?

$$\boxed{\frac{1}{n}\sum_{i=1}^{n}(u_1^T x_i - u_1^T \bar{x})^2}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(u_1^T (x_i - \bar{x})\right)^2$$

$$\frac{1}{n}\sum_{i=1}^{n} u_1^T (x_i - \bar{x})(x_i - \bar{x})^T u_1 \longrightarrow u_1^T \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T u_1 \longrightarrow \boxed{u_1^T S\, u_1}$$

We want to maximize this

So⋯

$$Max. \quad u_1^T S\, u_1$$

$$s.t: \quad u_1^T u_1 = 1$$

This is the difficult part. How would you solve it?

So⋯

$$Max.\quad u_1^T S\, u_1$$

$$s.t:\quad u_1^T u_1 = 1$$

This is the difficult part. How would you solve it?

Hint: Lagrange Multipliers

$$Max. \quad u_1^T S\, u_1$$

$$s.t: \quad u_1^T u_1 \; = 1$$

$$u_1^T S\, u_1 \; - \lambda\, (1 \; - u_1^T u_1)$$

To optimize, you take the derivative and equalize to zero.

$$Max. \quad u_1^T S\, u_1$$

$$s.t: \quad u_1^T u_1 \ = 1$$

$$u_1^T S\, u_1 \ - \lambda\,(1 \ - u_1^T u_1)$$

$$\frac{d}{du_i}\left(u_1^T S\, u_1 \ - \lambda\,(1 \ - u_1^T u_1)\right) = 0$$

$$Max. \quad u_1^T S\, u_1$$

$$s.t: \quad u_1^T u_1 \ = 1$$

$$u_1^T S\, u_1 \ - \lambda\,(1 \ - u_1^T u_1)$$

$$\frac{d}{du_i}\Big(u_1^T S\, u_1 \ - \lambda\,(1 \ - u_1^T u_1)\Big) = 0$$

$$2Su_i \ - \ \lambda 2u_i = 0$$

$$\boxed{Su_i \ = \lambda u_i}$$

$$Max. \quad u_1^T S\, u_1$$

$$s.t: \quad u_1^T u_1 \ = 1$$

$$u_1^T S\, u_1 \ - \lambda\, (1 \ - u_1^T u_1)$$

$$\frac{d}{du_i}\Big(u_1^T S\, u_1 \ - \lambda\, (1 \ - u_1^T u_1)\Big) = 0$$

$$2Su_i \ - \ \lambda 2u_i = 0$$

$$\boxed{Su_i \ = \lambda u_i}$$

What is this? Does it ring a bell?

$$Max. \quad u_1^T S \, u_1$$

$$s.t{:} \quad u_1^T u_1 \; = 1$$

$$u_1^T S \, u_1 \; - \lambda \, (1 \; - u_1^T u_1)$$

$$\frac{d}{du_i}\left(u_1^T S \, u_1 \; - \lambda \, (1 \; - u_1^T u_1)\right) = 0$$

$$2Su_1 \; - \; \lambda 2 u_1 = 0$$

$$\boxed{Su_1 \; = \lambda u_1}$$

$$u_1^T S u_1 \; = \lambda \, u_1^T u_1 \longrightarrow u_1^T S u_1 \; = \lambda$$

Which eigenvalue to use?

$$Max. \quad u_1^T S\, u_1$$

$$s.t: \quad u_1^T u_1 = 1$$

$$u_1^T S\, u_1 - \lambda\,(1 - u_1^T u_1)$$

$$\frac{d}{du_i}\Big(u_1^T S\, u_1 - \lambda\,(1 - u_1^T u_1)\Big) = 0$$

$$2Su_1 - \lambda 2u_1 = 0$$

$$\boxed{Su_1 = \lambda u_1}$$

$$u_1^T S u_1 = \lambda\, u_1^T u_1 \longrightarrow u_1^T S u_1 = \lambda$$

Which eigenvalue to use?
The highest, because it is the one that maximizes the variance of the projection.

$$Max.\quad u_1^T S\, u_1$$

$$s.t:\quad u_1^T u_1 = 1$$

$$u_1^T S\, u_1 - \lambda\,(1 - u_1^T u_1)$$

$$\frac{d}{du_i}\Big(u_1^T S\, u_1 - \lambda\,(1 - u_1^T u_1)\Big) = 0$$

$$2Su_1 - \lambda 2u_1 = 0$$

$$\boxed{Su_1 = \lambda u_1}$$

$$u_1^T S u_1 = \lambda\, u_1^T u_1 \longrightarrow u_1^T S u_1 = \lambda$$

https://quantitudepod.org/s3e03-principal-components-analysis-is-your-pal/

Go to the tutorial!

Thank you!

Orlando Sabogal-Cardona
@Antonio Sabogal
orlando.sabogal.20@ucl.ac.uk