

Mathématiques : CM

Contents

| | | |
|----------|---|-----------|
| 1 | Euclidean space - scalar product | 3 |
| 1.1 | Spectral Theorem | 9 |
| 1.2 | Theorem | 9 |
| 2 | Orthogonal Projection | 10 |
| 2.1 | Motivation | 10 |
| 2.2 | Definition | 11 |
| 2.3 | Theorem | 11 |
| 2.4 | Theorem | 11 |
| 2.5 | Property | 11 |
| 2.6 | Proof | 12 |
| 2.7 | Finding $P_F(u)$ | 13 |
| 3 | Data Analysis | 14 |
| 3.1 | Correlation coefficient | 18 |

1 Euclidean space - scalar product

Idea : Let E a vector space of finite dimension

The scalar product is defined as :

$$u \in E, v \in E, \langle u, v \rangle$$

$$\langle , \rangle = \begin{cases} E \times E \rightarrow \mathbb{R} \\ (u, v) \mapsto \langle u, v \rangle \end{cases}$$

When u and v are orthogonal (\perp) we have the following relation :

$$u \perp v \equiv \langle u, v \rangle = 0$$

We also have the following relation :

$$\|u\| = \sqrt{\langle u, u \rangle}$$

Standard scalar product of \mathbb{R}^3 :

$$u = (x_1, y_1, z_1)v = (x_2, y_2, z_2)$$

$$\begin{aligned} \langle u, v \rangle &= x_1x_2 + y_1y_2 + z_1z_2 \\ \implies \|u\| &= \sqrt{\langle u, u \rangle} = \sqrt{x_1^2 + y_1^2 + z_1^2} \end{aligned}$$

Definition : symmetric bilinear form

Let a mapping such that :

$$\phi : \begin{cases} E \times E \rightarrow \mathbb{R} \\ (u, v) \mapsto \phi(u, v) \end{cases}$$

ϕ is a symmetric bilinear form if :

1. It is bilinear : $\forall u_0 \in E, v \mapsto \phi(u_0, v)$ is a linear map
2. It is bilinear : $\forall v_0 \in E, u \mapsto \phi(u, v_0)$ is a linear map
3. It is symmetric : $\forall (u, v) \in E^2, \phi(u, v) = \phi(v, u)$ is a linear map

Explanation :

$$\phi = \begin{cases} E \times E \rightarrow \mathbb{R} \\ (u, v) \mapsto \phi(u, v) \end{cases}$$

Let $u_0 \in E$:

$$\text{The mapping : } \begin{cases} E \rightarrow \mathbb{R} \\ v \mapsto \phi(u_0, v) \end{cases} \text{ is linear}$$

$$\forall (v, v') \in E^2, \forall \alpha \in \mathbb{R}, \phi(u_0, \alpha v + v') = \alpha \phi(u_0, v) + \phi(u_0, v')$$

$$\text{If } v_0 \in E, \forall (u, u') \in E^2, \forall \alpha \in \mathbb{R}, \phi(\alpha u + u', v_0) = \alpha \phi(u, v_0) + \phi(u', v_0)$$

Example : standard scalar product of \mathbb{R}^3

$$u = (x_1, y_1, z_1)v = (x_2, y_2, z_2)$$

$$\phi(u, v) = x_1 x_2 + y_1 y_2 + z_1 z_2$$

If u is given (considered as a constant), then the mapping :

$v \mapsto x_1 x_2 + y_1 y_2 + z_1 z_2$ is a linear map.

Note that the coordinates of u are constant numbers.

Those of v are variables.

for example if $u = (1, 2, 3)$, then $v = (x_2, 2y_2, 3z_2)$

Similarly : if $v = (x_2, y_2, z_2)$ is given, then the mapping :

$u \mapsto x_1x_2 + y_1y_2 + z_1z_2$ is a linear map .

Note that the coordinates of v are being constant numbers.

Those of u are variables.

Furthermore, $\phi(u, v) = \phi(v, u)$

A symmetric bilinear form ϕ is positive definite if :

$$\forall u \in E, \phi(u, u) \geq 0 \wedge \phi(u, u) = 0 \implies u = 0_E$$

Example 1 : Standard scalar product of \mathbb{R}^3

$$\phi(u, u) = x_1^2 + y_1^2 + z_1^2 \geq 0$$

and

$$\phi(u, u) = 0 \implies \begin{cases} x_1^2 = 0 \\ y_1^2 = 0 \\ z_1^2 = 0 \end{cases} \implies u = 0_E$$

The standard scalar product is thus positive definite

Example 2 :

$\phi(u, v) = x_1x_2 + y_1y_2 - z_1z_2$ is a symmetric bilinear form

If $u = (x_1, y_1, z_1)$ is given (considered as a constant), then the mapping :

$v \mapsto x_1x_2 + y_1y_2 + z_1z_2$ is a linear map.

Note that the coordinates of u are constant numbers.

Those of v are the variables.

Thus the coordinates of u are the coefficient of the linear function.

Similarly, if v is given and constant, $u \mapsto \phi(u, v)$ is linear

$$\implies \phi(u, v) = \phi(v, u)$$

Positive definite ? No

$$\phi(u, u)x_1^2 + y_1^2 - z_1^2$$

If $u = (1, 0, 0)$, $\phi(u, u) = 1 > 0$

If $u = (0, 0, 1)$, $\phi(u, u) = -1 < 0$

Example 3 :

$$\phi(u, v) = x_1x_2 + y_1y_2 \quad \phi(u, u) = x_1^2 + y_1^2 \geq 0$$

But $\phi(u, u) = 0 \Rightarrow u = 0_E$ as for $u = (0, 0, 1)$, $\phi(u, u) = 0$, $u \neq 0_E$

A scalar product is a symmetric bilinear form, which is positive definite.

Then we note $\langle u, v \rangle$ instead of $\phi(u, v)$

We can define for $u \in E$,

$$\|u\| = \sqrt{\langle u, u \rangle} \implies \sqrt{\langle u, u \rangle} \in \mathbb{R} \text{ because } \langle u, u \rangle \geq 0$$

$$\|u\| = 0 \implies u = 0_E$$

Theorem : Matrix representation of a symmetric bilinear form :

Let $\mathcal{B} = (e_1, \dots, e_n)$ be a basis of E .

u has coordinates : $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ \dots \end{pmatrix}$ in \mathcal{B}

v has coordinates : $\begin{pmatrix} x_2 \\ y_2 \\ z_2 \\ \dots \end{pmatrix}$ in \mathcal{B}

Then : $\exists A \in \mathbb{M}_n(\mathbb{R})$, $\phi(u, v) = {}^T X_1 A X_2$ is symmetric

Examples :

Standard scalar product of \mathbb{R}^3 :

Example 1 :

\mathcal{B} standard basis of \mathbb{R}^3

$u = (x_1, y_1, z_1)$ has coordinates $X_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$

$v = (x_2, y_2, z_2)$ has coordinates $X_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$

${}^T X_1 X_2 = (x_1, y_1, z_1) \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$

$$= x_1 x_2 + y_1 y_2 + z_1 z_2 \implies A = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$${}^T X_1 A X_2 = x_1 x_2 + y_1 y_2 + z_1 z_2 = \langle u, v \rangle$$

Example 2 :

$$\phi(u, v) = x_1 x_2 + y_1 y_2 - z_1 z_2$$

$$\text{Then : } A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$${}^T X_1 A X_2 = (x_1, y_1, z_1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

$$= (x_1, y_1, z_1) \begin{pmatrix} x_2 \\ y_2 \\ -z_2 \end{pmatrix}$$

$$= x_1 x_2 + y_1 y_2 - z_1 z_2$$

Example 3 :

$$\phi(u, v) = x_1 x_2 + y_1 y_2$$

$$\text{Then : } A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$${}^T X_1 A X_2 = (x_1, y_1, z_1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

$$= (x_1, y_1, z_1) \begin{pmatrix} x_2 \\ y_2 \\ 0 \end{pmatrix}$$

$$= x_1 x_2 + y_1 y_2$$

1.1 Spectral Theorem

Spectral Theorem : all symmetric matrix A is diagonalizable.

We can choose an orthonormal eigenbasis $(\epsilon_1, \dots, \epsilon_n)$.

Then the transition matrix P satisfies to :

$${}^T P P = P \times {}^T P = I \Leftrightarrow P^{-1} = {}^T P$$

Note : if the eigenbasis is orthonormal, then :

$$\langle \epsilon_i, \epsilon_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

1.2 Theorem

Theorem : Let $E = \mathbb{R}^n$, let A be a symmetric matrix in $\mathcal{M}_n(\mathbb{R})$.

Let us define $\phi(u, v) = {}^T X_1 A X_2$.

ϕ is a scalar product \Leftrightarrow All eigenvalues of A are $\neq 0$.

Explanation in a an orthogonal eigenbasis \mathcal{B}' , $\phi(u, v) = {}^T X'_1 D X'_2$

Where D is a diagonal matrix.

2 Orthogonal Projection

Let V_0 be the orthogonal projection of V onto F .

V_0 is the element of F the closest to V .

2.1 Motivation

Assume you must find $\lambda \in \mathbb{R}$ such that :

$$v = \lambda u$$

The "best solution" is λ such that $v_0 = \lambda u$

Ex: $u = (1, 1), v = (2, 1)$

Solving $v = \lambda u$:
$$\begin{cases} \lambda = 2 \\ \lambda = 1 \end{cases}$$

$$\begin{aligned} v_0 &= \alpha u = (\alpha, \alpha) \\ (v - v_0) &\perp u \\ (2 - \alpha, 1 - \alpha) &\perp (1, 1) \\ \implies (2 - \alpha) \times 1 + (1 - \alpha) \times 1 &= 0 \\ \implies 3 - 2\alpha &= 0 \implies \alpha = \frac{3}{2} \\ \implies v_0 &= \frac{3}{2}(1, 1) = \left(\frac{3}{2}, \frac{3}{2}\right) \end{aligned}$$

The solution does not have a solution.

However, we can determine v_0 .

v_0 represent the solution with the samelless range of error.

2.2 Definition

Let $(E, <, >)$ be an euclidean space.

Let F be a linear subspace of E .

$$F^\perp = \{u \in E, \forall v \in F, \langle u, v \rangle = 0\}$$

2.3 Theorem

If F admits a basis $\mathcal{B} = (e_1, e_2, \dots, e_n)$, then $F^\perp = \mathcal{B}^\perp$

$$F^\perp = \{u \in E, \begin{cases} \langle u, e_1 \rangle = 0 \\ \langle u, e_2 \rangle = 0 \\ \dots \\ \langle u, e_n \rangle = 0 \end{cases}\}$$

2.4 Theorem

If F is finite-dimensional, then $F \oplus F^\perp = E$

$$\Leftrightarrow \forall u \in E, \exists! (v, w) \in F \times F^\perp, u = v + w$$

Then v is the orthogonal projection of u onto F .

2.5 Property

Let $u \in E$ and let $v_0 = P_F(u)$.

P_F is the orthogonal projection onto F .

v_0 is the element of F which minimizes $\|u - v\|^2$.

It is the solution of the optimization :

$$\text{Min}(\|u - v\|), \text{ constraint : } v \in F$$

2.6 Proof

Let $u = v_0 + w$ where $(v_0, w) \in F \times F^\perp$.

Then $\forall v \in F$:

$$\|u - v\|^2 = \langle u - v, u - v \rangle \implies u - v = (u - v_0) + (v_0 - v)$$

Indeed, $(u - v_0) = w \in F^\perp$ and $(v_0 - v) \in F$ as $(v_0, v) \in F^2$

$$\implies \|u - v\|^2 - \langle u - v, u - v \rangle = \langle (u - v_0) + (v_0 - v), (u - v_0) + (v_0 - v) \rangle$$

$$= \langle (u - v_0), (u - v_0) \rangle + 2 \langle (u - v_0), (v_0 - v) \rangle + \langle (v_0 - v), (v_0 - v) \rangle$$

$$= \|u - v_0\|^2 + \|v_0 - v\|^2 \geq \|u - v_0\|^2$$

$$P_F(u) \in F \Leftrightarrow \exists (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n, u = \lambda_1 e_1 + \dots + \lambda_n e_n$$

$$u - P_F(u) \in F^\perp \Leftrightarrow \begin{cases} \langle (u - \lambda_1 e_1 + \dots + \lambda_n e_n), e_1 \rangle = 0 \\ \dots \\ \langle (u - \lambda_1 e_1 + \dots + \lambda_n e_n), e_n \rangle = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \langle (\lambda_1 \langle e_1, e_1 \rangle + \dots + \lambda_n \langle e_n, e_n \rangle), e_1 \rangle = \langle u, e_1 \rangle \\ \dots \\ \langle (\lambda_1 \langle e_1, e_1 \rangle + \dots + \lambda_n \langle e_n, e_n \rangle), e_n \rangle = \langle u, e_n \rangle \end{cases}$$

2.7 Finding $P_F(u)$

Given F admits a basis $\mathcal{B}_F = (\epsilon_1, \dots, \epsilon_n)$

$$\text{We write } P_F(u) \in F \implies u - P_F(u) \in F^\perp$$

$$\implies u = P_F(u) + (u - P_F(u))$$

$$P_F(u) \in F \Leftrightarrow \exists(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n, P_F(u) = \alpha_1\epsilon_1 + \dots + \alpha_n\epsilon_n$$

$$u - P_F(u) \in F^\perp \Leftrightarrow \begin{cases} < u - P_F(u), \epsilon_1 > = 0 \\ \dots \\ < u - P_F(u), \epsilon_n > = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} < u - \alpha_1\epsilon_1 + \dots + \alpha_n\epsilon_n, \epsilon_1 > = 0 \\ \dots \\ < u - \alpha_1\epsilon_1 + \dots + \alpha_n\epsilon_n, \epsilon_n > = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} < \alpha_1 < \epsilon_1, \epsilon_1 > + \dots + \alpha_n < \epsilon_n, \epsilon_n >, \epsilon_1 > = < u, \epsilon_1 > \\ \dots \\ < \alpha_1 < \epsilon_1, \epsilon_1 > + \dots + \alpha_n < \epsilon_n, \epsilon_n >, \epsilon_n > = < u, \epsilon_n > \end{cases}$$

3 Data Analysis

Example of a dataset :

| Student | Class | Average | Sport | French | Math | Physics | English |
|----------|-------|---------|-------|--------|-------|---------|---------|
| Student1 | A1 | 1 | SP1 | FR1 | MATH1 | PHY1 | ENG1 |
| Student2 | B2 | 1 | SP2 | FR2 | MATH2 | PHY2 | ENG2 |
| Student3 | C1 | 1 | SP3 | FR3 | MATH3 | PHY3 | ENG3 |
| Student4 | B2 | 1 | SP4 | FR4 | MATH4 | PHY4 | ENG4 |

Table 1: Table of marks average in a high school.

The question that we address:

- Consider the student good in maths. Are they good in physics ?

We try to find a relation :

$$PHY = \alpha MAT + \beta + Err_1$$

And compare the error with :

$$PHY = \beta + Err_2$$

The "best" β is the Average of PHY column.

Is $\|Err_1\| < \|Err_2\|$?

If yes, is $\alpha > 0$?

For the model :

$$PHY = \alpha MAT + \beta + Err_1$$

we look for the best solution of :

$$(S) : \begin{cases} PHY_1 = \alpha MAT_1 + \beta \\ PHY_2 = \alpha MAT_2 + \beta \\ \dots \end{cases}$$

$$\Leftrightarrow PHY = \alpha MAT + \beta \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}$$

Where PHY and MAT are the whole column.

Optimal values : α_0 and β_0 such that $\alpha_0 MAT + \beta_0 \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}$ is the orthogonal

projection of PHY on $Span(MAT, \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix})$

- The students good at maths and sport, do they tend to be even better than those good only at maths ?

$$PHY = \alpha MAT + \beta \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} + \gamma SP + Err_3$$

Find $\alpha_0, \beta_0, \gamma_0$ optimal :

Is $\|Err_3\| < \|Err_1\|$?

If yes, sign of β ?

Let X be a column of the table.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

X is an identically distributed random variable and independant.

Empirical expectation of X :

$$\hat{\mathbb{E}}(X) = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\begin{matrix} ^T \\ \left(\begin{matrix} 1 \\ \dots \\ 1 \end{matrix} \right) \cdot X \end{matrix}}{n}$$

Empirical variance of X :

$$\widehat{\text{Var}}(X) = \hat{\mathbb{E}}((X - \hat{\mathbb{E}}(X))^2)$$

$$= \frac{(x_1 - \hat{\mathbb{E}}(X))^2 + (x_2 - \hat{\mathbb{E}}(X))^2 + \dots + (x_n - \hat{\mathbb{E}}(X))^2}{n}$$

$$\text{Let } X_c = X - \hat{\mathbb{E}}(X) \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix}$$

$$X_c = \begin{pmatrix} x_1 - \hat{\mathbb{E}}(X) \\ x_2 - \hat{\mathbb{E}}(X) \\ \dots x_n - \hat{\mathbb{E}}(X) \end{pmatrix}$$

X_c is the centered version of X.

$$\text{Thus } \widehat{\text{Var}} = \frac{(x_1 - \hat{\mathbb{E}}(X))^2 + (x_2 - \hat{\mathbb{E}}(X))^2 + \dots + (x_n - \hat{\mathbb{E}}(X))^2}{n}$$

$$= \frac{1}{n} \begin{pmatrix} x_1 - \hat{\mathbb{E}}(X) & x_2 - \hat{\mathbb{E}}(X) & \dots & x_n - \hat{\mathbb{E}}(X) \end{pmatrix} \begin{pmatrix} x_1 - \hat{\mathbb{E}}(X) \\ x_2 - \hat{\mathbb{E}}(X) \\ \dots \\ x_n - \hat{\mathbb{E}}(X) \end{pmatrix}$$

$$= \frac{^T X_c X_c}{n} = \frac{\|X_c\|^2}{n}$$

Empirical standard deviation of X :

$$\hat{\sigma}(X) = \sqrt{\widehat{\text{Var}}(X)} = \sqrt{\frac{^T X_c X_c}{n}} = \frac{\|X_c\|}{\sqrt{n}}$$

Consider two columns $X = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ and $Y = \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}$

Empirical covariance of X and Y :

$$\begin{aligned} \widehat{\text{Cov}}(X, Y) &= \hat{\mathbb{E}}((X - \hat{\mathbb{E}}(X)) \times (Y - \hat{\mathbb{E}}(Y))) \\ &= \frac{1}{n} (x_1 - \hat{\mathbb{E}}(X))(y_1 - \hat{\mathbb{E}}(Y)) + \dots + (x_n - \hat{\mathbb{E}}(X))(y_n - \hat{\mathbb{E}}(Y)) \\ &= \frac{1}{n} \begin{pmatrix} x_1 - \hat{\mathbb{E}}(X) & x_2 - \hat{\mathbb{E}}(X) & \dots & x_n - \hat{\mathbb{E}}(X) \end{pmatrix} \begin{pmatrix} y_1 - \hat{\mathbb{E}}(Y) \\ y_2 - \hat{\mathbb{E}}(Y) \\ \dots \\ y_n - \hat{\mathbb{E}}(Y) \end{pmatrix} \\ &= \frac{^T X_c Y_c}{n} \\ &= \frac{1}{n} \langle X_c, Y_c \rangle \end{aligned}$$

Empirical correlation coefficient of X and Y :

$$\hat{c}(X, Y) = \frac{\widehat{\text{Cov}}(X, Y)}{\hat{\sigma}(X)\hat{\sigma}(Y)}$$

Note that $\widehat{\text{Cov}}(X, X) = \widehat{\text{Var}}(X)$ and $\widehat{\text{Cov}}(Y, Y) = \widehat{\text{Var}}(Y)$

Let $D = \text{Dataset} = (X \ Y) = \begin{pmatrix} x_1 & y_1 \\ \dots & \dots \\ x_n & y_n \end{pmatrix}$

$D_c = \text{centered data set} = (X_c \ Y_c)$

The Empirical covariance matrix of D is :

$$\widehat{\text{Cov}}(D) = \begin{pmatrix} \widehat{\text{Var}}(X) & \widehat{\text{Cov}}(X, Y) \\ \widehat{\text{Cov}}(X, Y) & \widehat{\text{Var}}(Y) \end{pmatrix} = \frac{^T D_c D_c}{n}$$

$$^T D_c D_c = \begin{pmatrix} x_1 - \hat{\mathbb{E}}(X) & \dots & x_n - \hat{\mathbb{E}}(X) \\ y_1 - \hat{\mathbb{E}}(Y) & \dots & y_n - \hat{\mathbb{E}}(Y) \end{pmatrix} \begin{pmatrix} x_1 - \hat{\mathbb{E}}(X) & y_1 - \hat{\mathbb{E}}(Y) \\ \dots & \dots \\ x_n - \hat{\mathbb{E}}(X) & y_n - \hat{\mathbb{E}}(Y) \end{pmatrix}$$

$$= \begin{pmatrix} {}^T X_c X_c & {}^T X_c Y_c \\ {}^T Y_c X_c & {}^T Y_c Y_c \end{pmatrix}$$

Ex : find $\widehat{\text{Var}}(aX + bY)$ where $aX + bY = D \begin{pmatrix} a \\ b \end{pmatrix}$

$$D \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \dots & \dots \\ x_n & y_n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ax_1 + by_1 \\ ax_2 + ny_2 \\ \dots \\ ax_n + by_n \end{pmatrix} = aX + bY$$

$$\widehat{\text{Var}}(X) = {}^T X_c X_c$$

$$\begin{aligned} \widehat{\text{Var}}(D \begin{pmatrix} a \\ b \end{pmatrix}) &= (D_c \begin{pmatrix} a \\ b \end{pmatrix})(D_c \begin{pmatrix} a \\ b \end{pmatrix}) \times \frac{1}{n} \\ &= (a \ b) {}^T D_c D_c \begin{pmatrix} a \\ b \end{pmatrix} \times \frac{1}{n} \\ &= (a \ b) \widehat{\text{Cov}}(D) \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

3.1 Correlation coefficient

Definition :

$$\hat{r}(X, Y) = \frac{\widehat{\text{Cov}}(X, Y)}{\hat{\sigma}(X)\hat{\sigma}(Y)}$$

$$\widehat{\text{Cov}}(X, Y) = \frac{1}{n} {}^T X_c Y_c = \frac{1}{n} \langle X_c, Y_c \rangle$$

$$\hat{\sigma}(X) = \sqrt{\frac{1}{n} {}^T X_c X_c}$$

$$= \frac{1}{\sqrt{n}} \sqrt{^T X_c X_c} = \frac{1}{\sqrt{n}} \|X_c\|$$

$$\hat{\sigma}(Y) = \frac{1}{\sqrt{n}} \|Y_c\|$$

$$\implies \hat{r}(X, Y) = \frac{\langle X_c, Y_c \rangle}{\|X_c\| \cdot \|Y_c\|}$$

In all euclidean space,

$$-\|u\| \cdot \|v\| \leq \langle u, v \rangle \leq \|u\| \cdot \|v\|$$

Explanation :

$$\begin{cases} u = u_1 + u_2 \\ u_1 = \lambda V \\ \langle u_2, v \rangle = 0 \\ \lambda \in \mathbb{R} \end{cases}$$

$$\begin{aligned} \langle u, v \rangle &= \langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle \\ &= \lambda \langle v, v \rangle = \lambda \|v\|^2 \end{aligned}$$

$$\lambda \|v\|^2 = \lambda \|v\| + \|v\|$$

$$\implies -\|u\|^2 \leq \lambda \|v\| \leq \|u\|$$

$$\begin{aligned} (\lambda \|v\|)^2 &= \lambda^2 \langle v, v \rangle \\ \langle \lambda v, \lambda v \rangle &= \langle u_1, u_1 \rangle \end{aligned}$$

$$= \|u_1\|^2 \leq \|u\|^2$$

Because $\|u\|^2 = \|u_1\|^2 + \|u_2\|^2$

Consequence : $-1 \leq \hat{r}(X, Y) \leq 1$

$$\widehat{\text{Cov}}(aX + bY, a'X + b'Y) = (a \quad b) \widehat{\text{Cov}}(D) \begin{pmatrix} a' \\ b' \end{pmatrix}$$