In this section we assume that we are working with a data matrix

$$A = [\; X_1 \quad X_2 \quad \dots \quad X_M \;]$$

which has been demeaned. That is $m_{X_i} = 0$, or equivalently $X_i = \widetilde{X}_i$ for $i = 1, \dots, M$.

Example.

A data matrix with demeaned exam scores:

$$A = \begin{array}{r} \\ \text{Aly} \\ \text{Bob} \\ \text{Chen} \\ \text{Deb} \\ \text{Emma} \\ \text{Finn} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Ex 1} & \text{Ex 2} & \text{Ex 3} \end{array} \\ \begin{bmatrix} -24 & 1 & -40 \\ -3 & -2 & -6 \\ 29 & 5 & 17 \\ 26 & -2 & 9 \\ -43 & 5 & 30 \\ 15 & -7 & -10 \end{bmatrix} \end{array}$$

Definition

Let $A = \begin{bmatrix} X_1 & \dots & X_M \end{bmatrix}$ be a demeaned data matrix.

- The *1st principal axis* of A is a vector

$$\mathbf{u}_1 = \begin{bmatrix} c_1 \\ \vdots \\ c_M \end{bmatrix}$$

such that $||\mathbf{u}_1|| = 1$ and the variance of the vector

$$Y_1 = A\mathbf{u}_1 = c_1 X_1 + \ldots + c_M X_M$$

is the largest possible.

- The vector Y_1 is called the *1st principal component* of A.

<u>Computations of u_1</u>

Take $u = \begin{bmatrix} c_1 \\ \vdots \\ c_M \end{bmatrix}$, $\|u\| = 1$, and let $Y = A \cdot u = c_1 X_1 + \ldots + c_M X_M$

<u>Check</u>: Since $X_i = \tilde{X}_i$ for $i = 1, \ldots, M$ thus we have $Y = \tilde{Y}$.

This gives : $\mathrm{Var}(Y) = \frac{1}{N} Y^T Y$

$$= \frac{1}{N}(Au)^T(Au)$$

$$= \frac{1}{N} u^T A^T A u$$

$$= u^T \underbrace{\left(\tfrac{1}{N} A^T A\right)}_{C_A} u$$

the quadratic form defined by the matrix C_A

C_A - the covariance matrix of A (since $A = \tilde{A}$)

<u>Constrained optimization of quadratic forms gives:</u>

1) $\mathrm{Var}(Y)$ is the largest if u is a unit eigenvector corresponding to the largest eigenvalue λ_1 of C_A.
2) For this choice of u we have $\mathrm{Var}(Y) = \lambda_1$.

Proposition

Given a demeaned data matrix $A = \begin{bmatrix} X_1 & \dots & X_M \end{bmatrix}$ the 1st principal axis of A is a vector

$$\mathbf{u}_1 = \begin{bmatrix} c_1 \\ \vdots \\ c_M \end{bmatrix}$$

such that $||\mathbf{u}_1|| = 1$ and $\mathbf{u}_1$ is an eigenvector of the covariance matrix C_A corresponding to the largest eigenvalue of this matrix.

Moreover, if $Y_1 = A\mathbf{u}_1$ is the 1st principal component of A then $\mathrm{Var}(Y_1) = \lambda_1$ where λ_1 is the largest eigenvalue of the covariance matrix C_A.

Note:

$$\mathrm{Var}(A) = \mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_M) = \mathrm{tr}\, C_A = \underbrace{\lambda_1 + \lambda_2 + \dots + \lambda_M}_{\substack{\text{all eigenvalues of } C_A \\ (\lambda_i \geq 0 \text{ since } C_A \\ \text{is positive semidefinite})}}$$

This gives: $0 \leq \mathrm{Var}(Y_1) = \lambda_1 \leq \lambda_1 + \dots + \lambda_M = \mathrm{Var}(A).$