

$$S_{ij} = \text{cov}(f_i, f_j)$$

$i: 1 \rightarrow d$
 $j: 1 \rightarrow d$

$$\boxed{\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)}$$

$$\text{Cov}(f_i, f_i) = \text{Var}(f_i)$$

$$\left\{ \begin{array}{l} \checkmark \text{Cov}(x, x) = \text{Var}(x) \text{ --- (1)} \\ \checkmark \text{Cov}(f_i, f_j) = \text{Cov}(f_j, f_i) \text{ --- (2)} \end{array} \right.$$

Variance of
features

$$S = \begin{bmatrix} s_{11} & s_{12} & s_{13} & \dots & s_{1d} \\ s_{21} & s_{22} & s_{23} & \dots & s_{2d} \\ s_{31} & s_{32} & s_{33} & \dots & s_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{d1} & s_{d2} & s_{d3} & \dots & s_{dd} \end{bmatrix}$$

s_{ij} s_{ji}

$d \times d$

Symmetric matrix

$$\begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 3 & 2 & 5 \end{bmatrix} \end{matrix}$$

3×3

$$A_{32} = S = A_{23}$$

$$A_{21} = A_{12}$$

$$\boxed{A_{ij} = A_{ji} \quad \forall i, j}$$

$$\text{Cov}(f_1, f_2) = \frac{1}{n} \sum_{i=1}^n x_{i1} * x_{i2}$$

Diagram illustrating the data matrix X with columns f_1 and f_2 . The matrix is shown as a vertical stack of rows, with the first two columns labeled f_1 and f_2 . The rows are indexed 1 to n . The elements x_{i1} and x_{i2} are highlighted in the first two columns. A bracket indicates the dot product $f_1 \cdot f_2$ for each row i .

$$\text{Cov}(f_1, f_2) = \frac{1}{n} (f_1^T f_2)$$

we'll assume S to be this matrix

$$S_{d \times d} = \frac{1}{n} (X^T)_{d \times n} (X)_{n \times d} = \text{data-matrix}$$

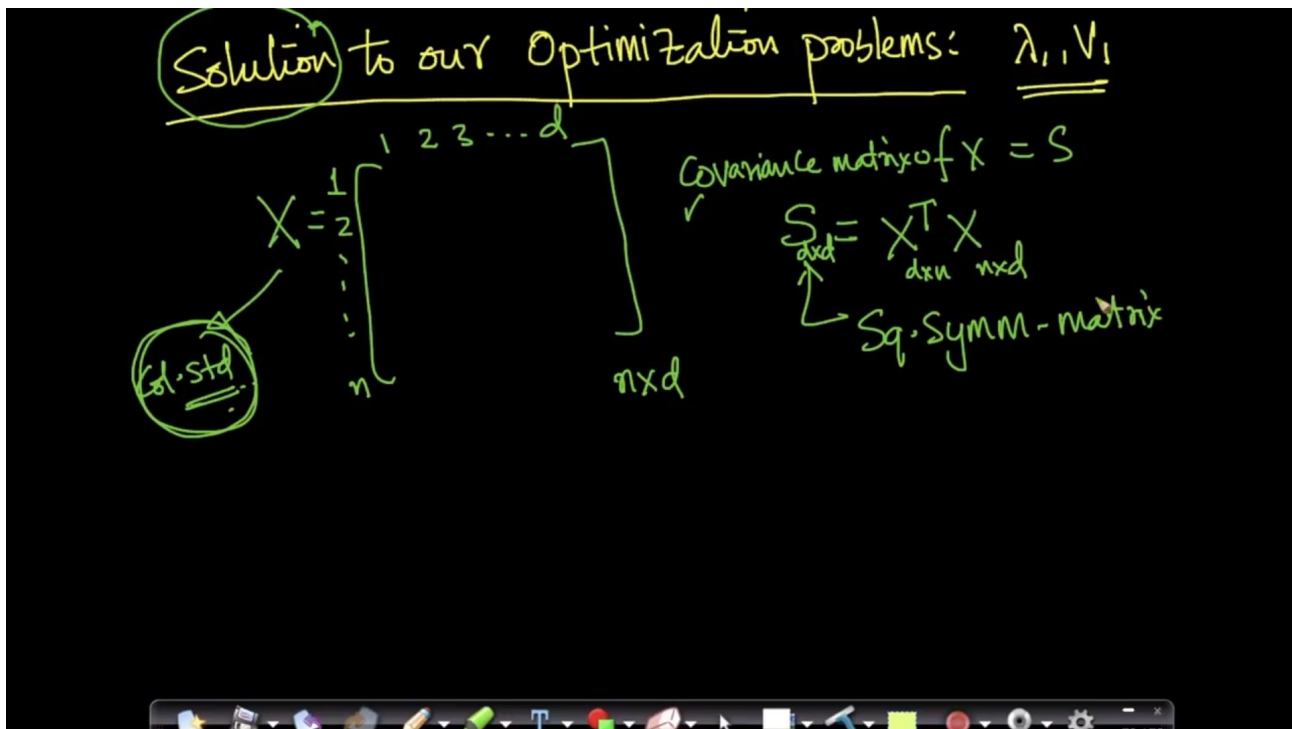
(*) assuming X has been col. std

(LHS) $S_{ij} = \text{Cov}(f_i, f_j) = \frac{f_i^T f_j}{n}$

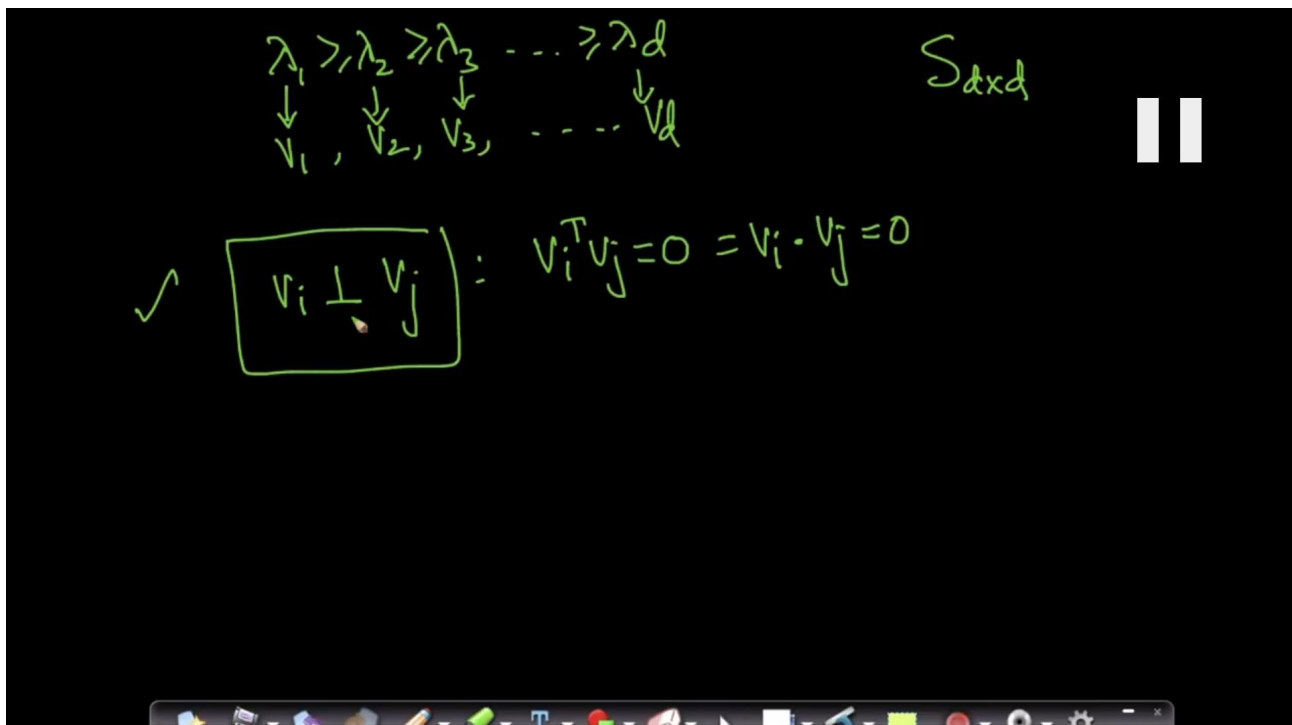
$$S_{d \times d} = \frac{1}{n} X_{d \times n}^T X_{n \times d}$$

if X has been col. std

this X transpose * X needs to be divided by n here



if we solve the optimisation problem of PCA, any of the two,
since it's a $d \times d$ matrix you will get d eigenvalues



u_1 which we want to find in PCA, is equal to v_1 , which is the biggest eigen vector

each eigen vector has it's own variance
 v_1 has max variance

