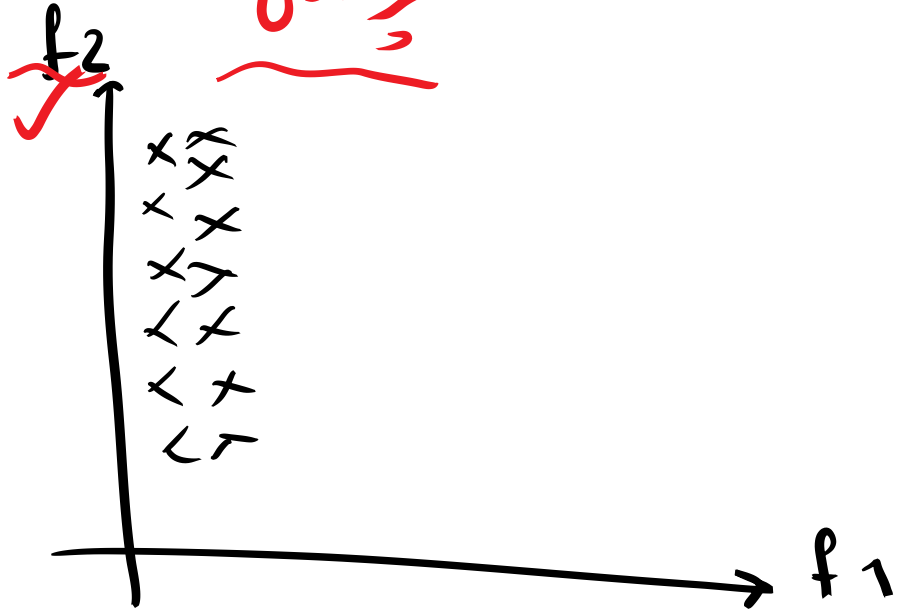
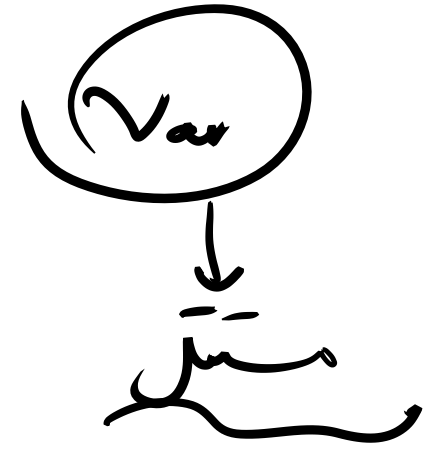
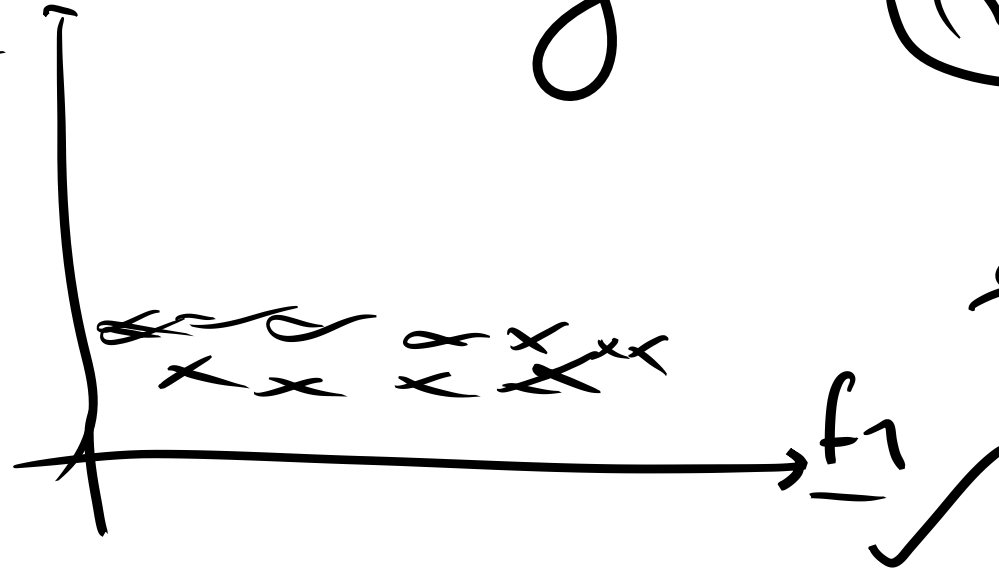


داده‌های  
تستی



feature Engineering

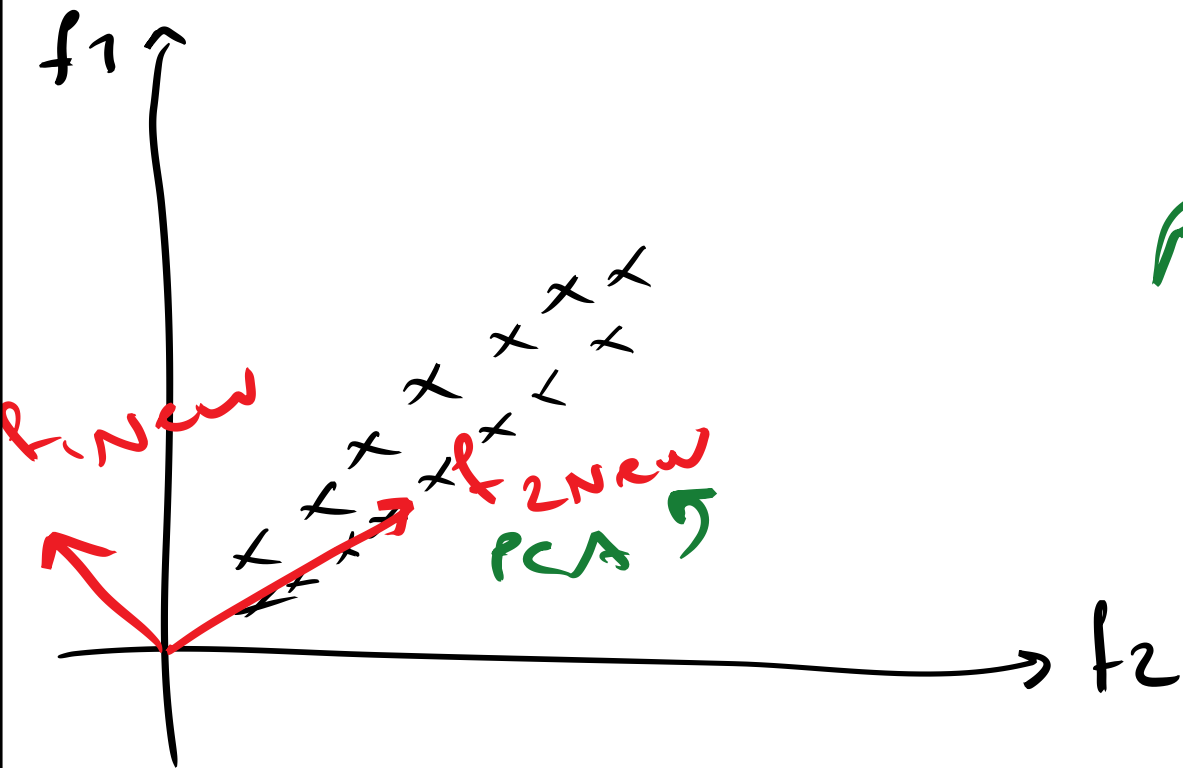


$$Cov = \begin{pmatrix} Var(f_1) & C_{1,2} & C_{1,3} & C_{1,4} \\ C_{2,1} & V_2 & C_{2,3} & C_{2,4} \\ C_{3,1} & C_{3,2} & V_3 & C_{3,4} \\ C_{4,1} & C_{4,2} & C_{4,3} & V_4 \end{pmatrix}$$

→ Feature Selection

① مارکس: همبستگی - Diagonal

② دیگر عناصر: نشان می‌دهد که  $Var$  بسیار کوچک باشد.



PCA  
Principal Component Analysis

تحلیل مؤلفه‌های اصلی

محاسبه جهت و بزرگای  
معمولی

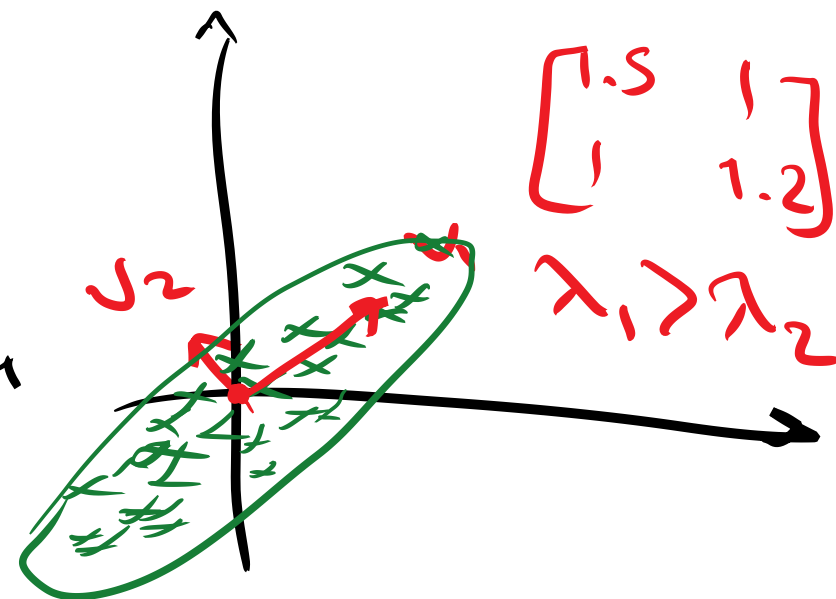
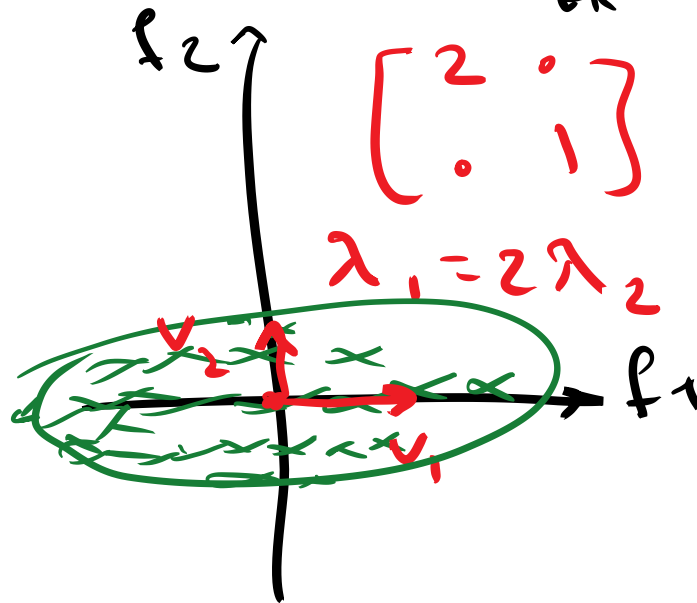
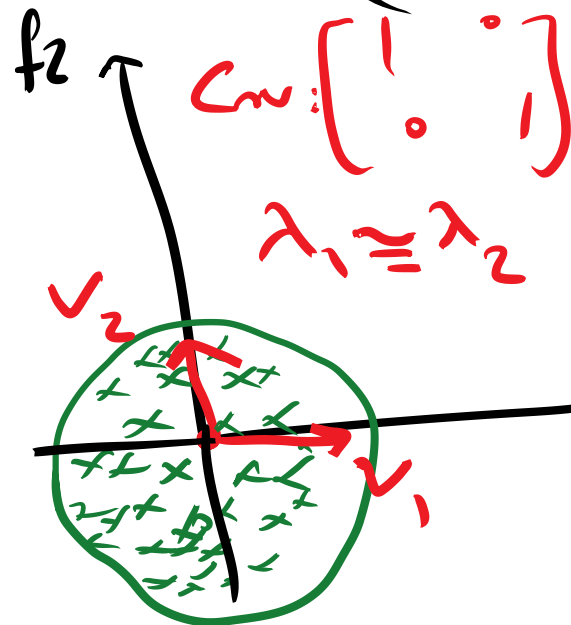
unsupervised

$A_{k \times k}$

$$\det(A - \lambda I) = 0 \quad \lambda_i \rightarrow \underline{\sigma_i}, \text{ and } \mathbb{K}$$

$$(A - \lambda_i I) V_i = 0$$

$$V_i \xrightarrow{\mathbb{K}} \underline{\sigma_i}, \text{ and } \mathbb{K} \quad V_{1 \times k}$$



$$X_{n \times f}$$

n تعداد داده ها

f ویژگی ها

عدد k

$$1) X = X - \mu \quad \text{Zero mean}$$

$$2) C = \frac{(X - \mu)^T (X - \mu)}{n - 1} \quad \text{Cov}$$

3)

$\lambda_1 \dots \lambda_f$   
 $v_1 \dots v_f$

حاصل برداشتن  
مهمترین  
 $v_i$  (f تا 1)

desc.  
4) Sort ( $\lambda_i$ )  $\rightarrow v_i$

$\lambda_1 \dots \lambda_f$   
 $v_1 \dots v_k$

5)

$$W_{f \times k} \\ X_{Zero} \times W = M_{n \times k}$$

$$\sum_{i=1}^k \lambda_i$$

$\lambda_i \rightarrow \begin{matrix} \text{desc.} \\ \text{Sorted} \end{matrix}$

$PCA(5)$   
 $\uparrow$   
 $k$

---


$$\sum_{i=1}^k \lambda_i$$

$> .9$   
 $.95$

$P(A(.95))$   
 $\leftarrow$

$k$   
 $\leftarrow$

