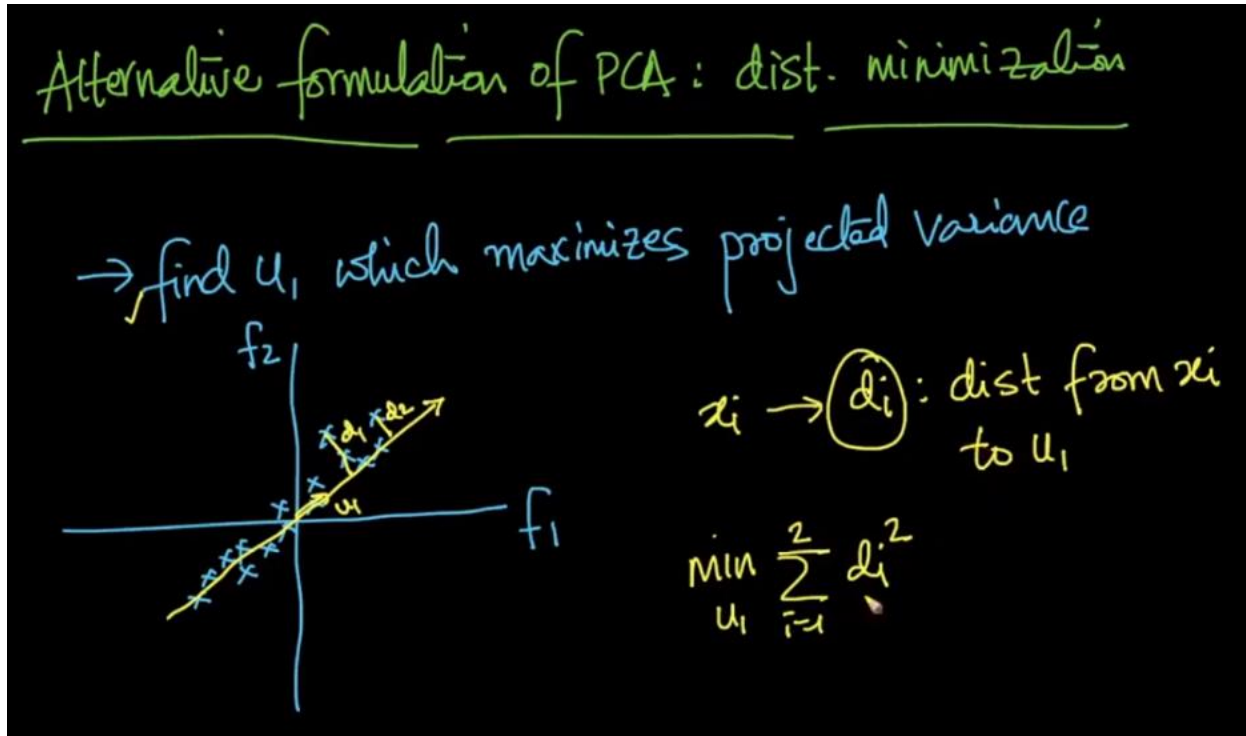


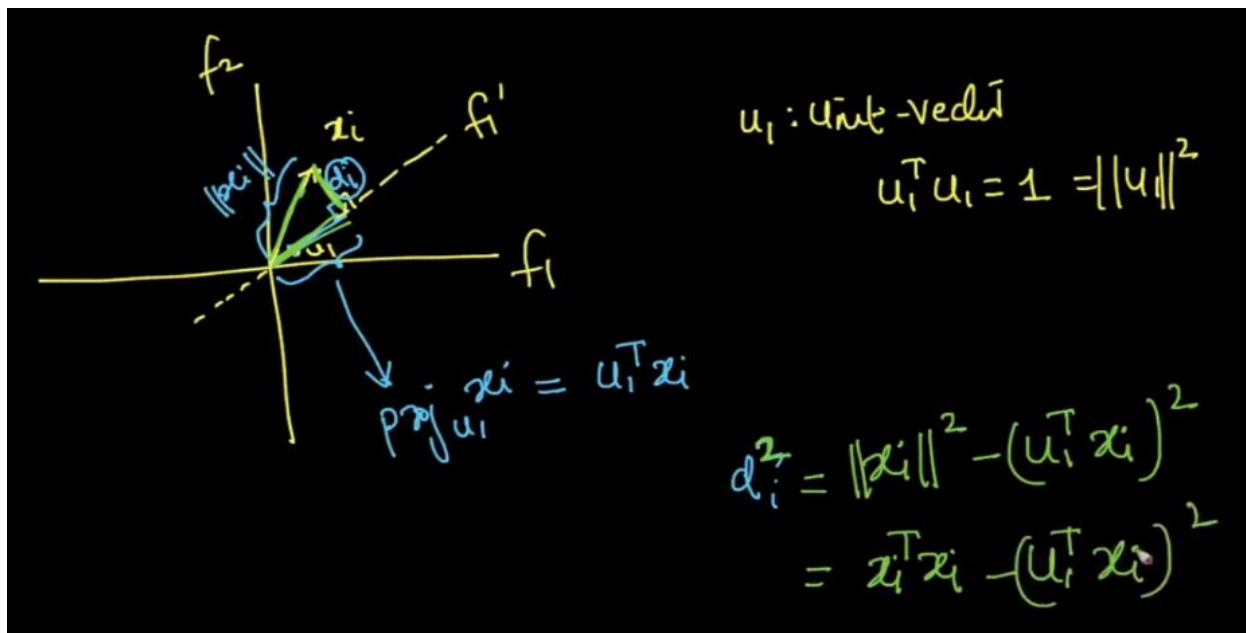
Earlier we have seen for a technique finding a new feature or direction which was finding the direction where the variance was maximum, in this we will see a new technique called **distance minimization**.

**Objective:** In this technique the objective is to find the direction from which the distance ( $d_i$ ) is minimum from ( $x_i$ ).



How we find distance of  $x_i$  and  $u_1$ . Since the projection of  $x_i$  on  $u_1$  forms a right angle triangle so we can apply Pythagoras to find  $d_i$ .

The calculation is given in below fig.



Now our ultimate aim is to find a direction  $u_1$  such that it has minimum distance from  $x_i$ , given a constraint that  $u^T u = 1$

Handwritten diagram illustrating the distance minimization problem for PCA. The objective is to minimize  $u_1$  (circled in green) subject to the constraint  $u^T u = 1$  (circled in orange). The distance squared  $d_i^2$  is shown as  $\|x_i\|^2 - (u_1^T x_i)^2$ , where  $\|x_i\|^2$  is circled in green and  $(u_1^T x_i)^2$  is circled in blue. A matrix  $X$  is shown with rows  $x_i^T$  (indicated by arrows).

We've seen two different techniques for PCA, one is distance minimization PCA and other is variance maximization PCA, and eventually both lead to the same  $u_1$ .

Handwritten diagram comparing distance minimization PCA and variance maximization PCA. The distance minimization problem (labeled "dist min PCA") is shown with the objective  $\min u_1$  (circled in green) and constraint  $u^T u = 1$  (circled in orange). The variance maximization problem (labeled "Variance maximization PCA") is shown with the objective  $\max u_1$  (circled in green) and constraint  $u^T u = 1$  (circled in orange). The variance maximization formula is  $\frac{1}{n} \sum_{i=1}^n (u_1^T x_i)^2$ , where  $(u_1^T x_i)^2$  is circled in green. A matrix  $X$  is shown with rows  $x_i^T$  (indicated by arrows).