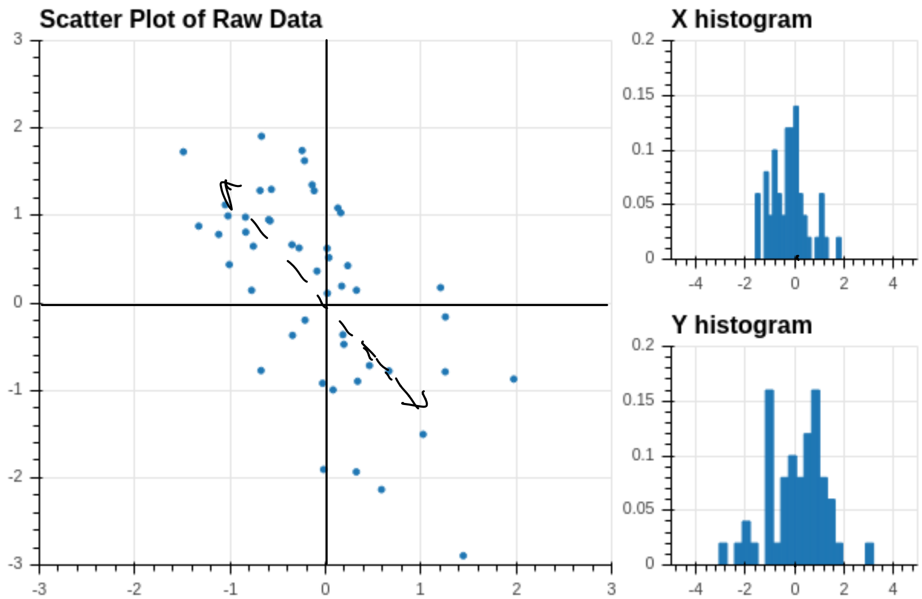
Geometry of Scores

A look at sample data



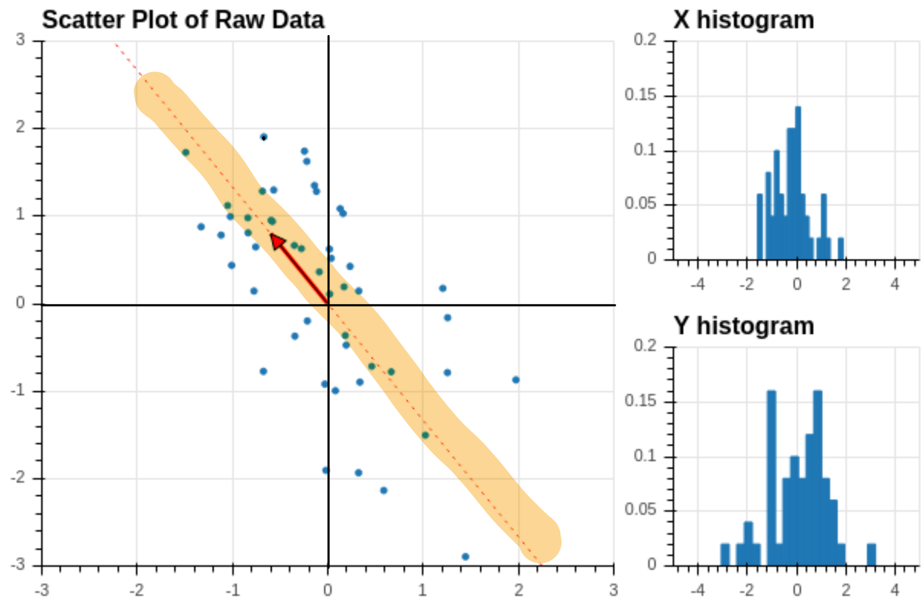
50 points in $\mathbb{R}^2$

centered set of data

$N = 50 \quad K = 2$

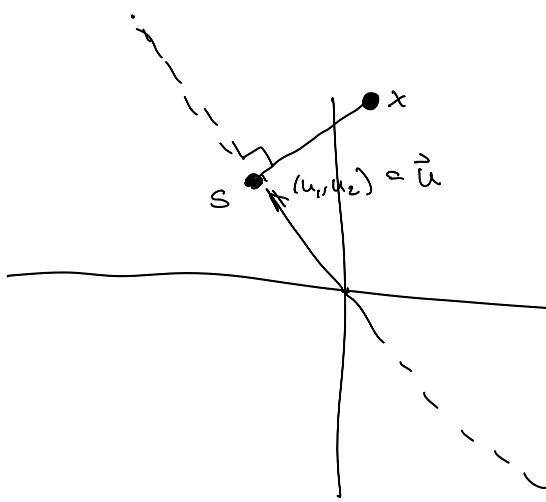
$$X_0 = \begin{array}{c c} & 2 \\ 50 & \begin{pmatrix} \vdots & \vdots \\ \vdots & \vdots \end{pmatrix} \end{array}$$

A direction in the data (a score)



$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is a "feature space" $\mathbb{R}^2$

points in an "interesting" direction.



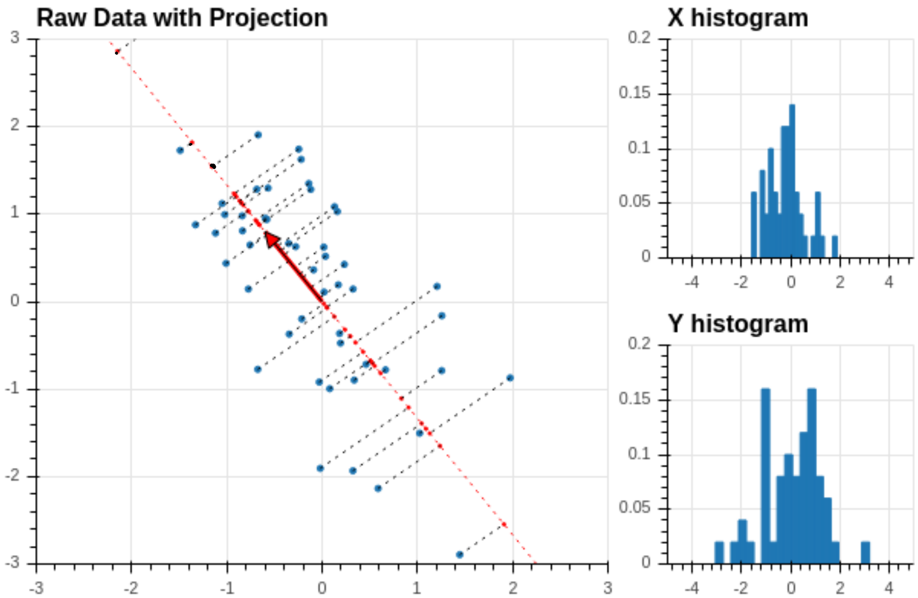
$(x \cdot \vec{u})\vec{u}$ = projection of x onto the line in u direction.

$\|u\| = 1$

Value $x \cdot \vec{u}$ is a score

$x_1 u_1 + x_2 u_2$

Projection onto the score direction



$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad \|u\|^2 = 1$$

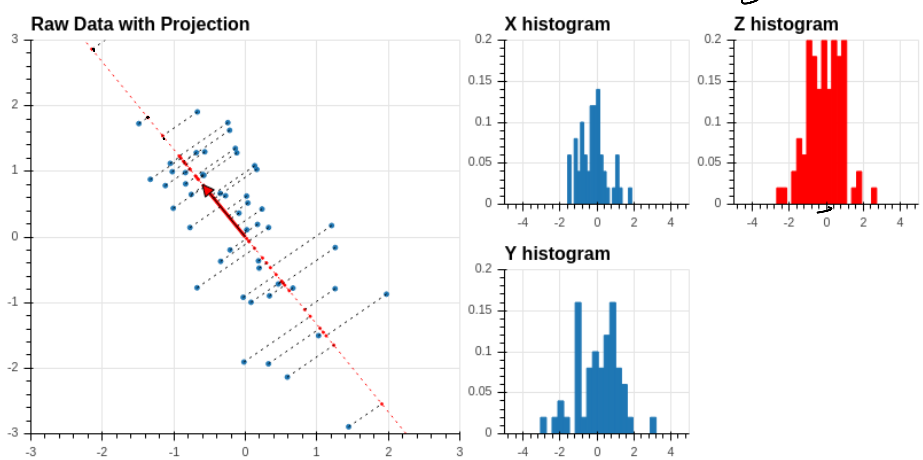
$$S = X_0 u$$

entries of S are dot products of rows of X_0 – pts (x_i, y_i) – with vector u

= orthogonal projections of data onto direction $\vec{u}$.

Histogram of the score

scores $(X_0 \cdot u)$



Variance of score $= u^T D_0 u$ as $\vec{u}$ varies.