Elaboration of Projection view of Linear Regression

Lecture 3 (Feb 2)
David Rapkes

We can also gain a geometric interpretation on linear regression.

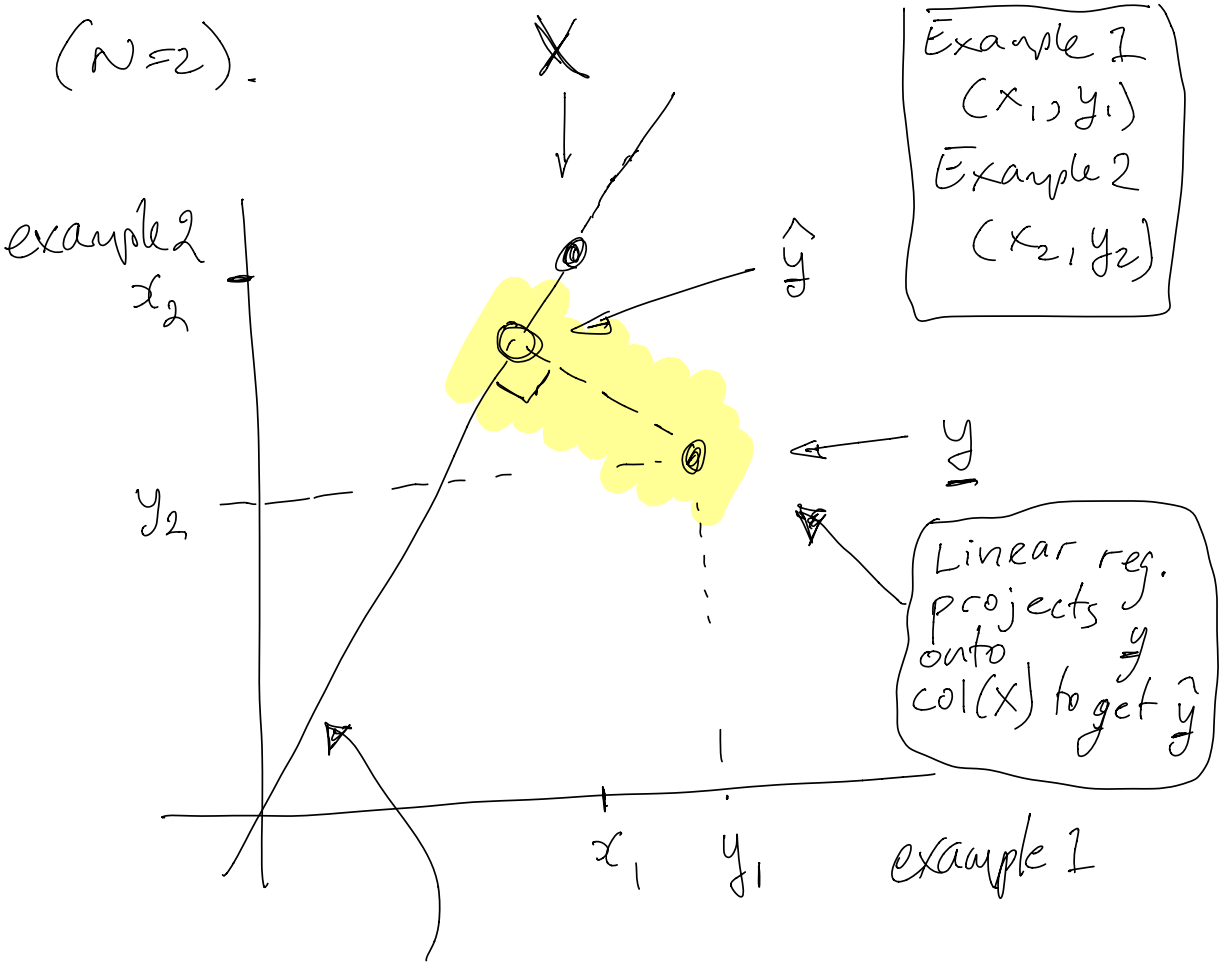
Recall $\hat{y} = \underline{x}^T (X^T X)^{-1} X^T \underline{y}$, and write

$\hat{\underline{y}} = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_N \end{bmatrix}$. With this, we have

$$\boxed{\hat{\underline{y}} = X(X^T X)^{-1} X^T \underline{y} = P\underline{y}}$$

P is a projection of $\underline{y}$ onto the column space of design matrix X. Often called the "hat matrix" because it "turns $\underline{y}$ into $\hat{\underline{y}}$."

For a simple example, suppose we have 1 feature ($D=1$) and 2 examples ($N=2$).



line corresponds to column space of X

(here, 1D features and so just points $S: \{\alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : \alpha \in \mathbb{R}\}$