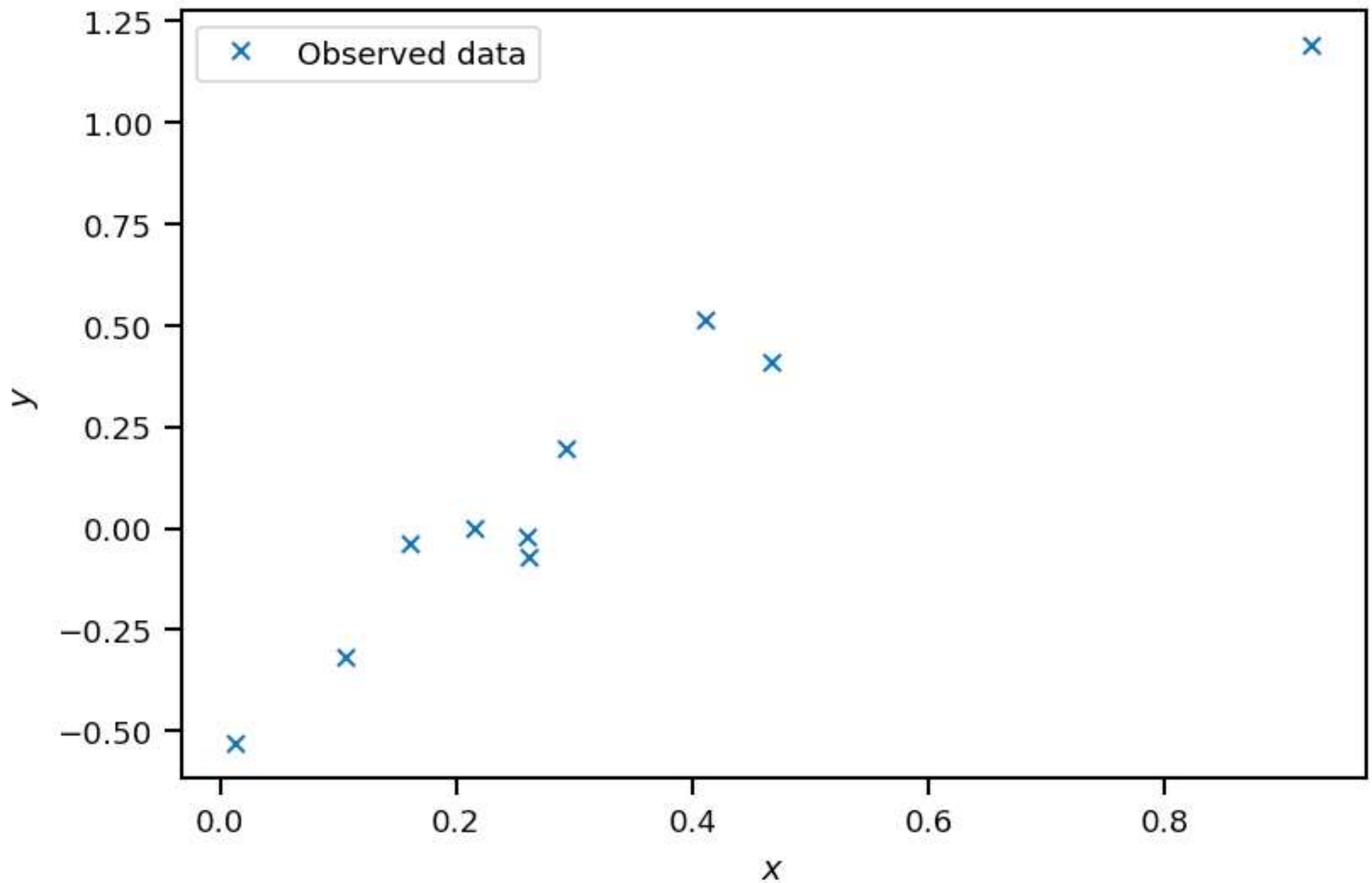
Lecture 13: Linear Regression via Least Squares

Professor Ilias Bilionis

Linear regression with a single variable

PREDICTIVE
SCIENCE LABORATORY

Synthetic example



PREDICTIVE
SCIENCE LABORATORY

Regression model

$$y = w_0 + w_1 x$$

weights
(parameters)

$\underline{w} = (w_0, w_1)$: weight vector

$\underline{w} = ?$ using data

PREDICTIVE
SCIENCE LABORATORY

Least squares loss function

function of the parameters

$$L(\underline{w}) := \sum_{i=1}^{n} \left[y_i - \underbrace{(w_0 + w_1 x_i)}_{\text{prediction}} \right]^2$$

Find $\underline{w}$ by $\min L(\underline{w})$.

$$\underline{w}^* = \arg\min L(\underline{w}).$$

PREDICTIVE
SCIENCE LABORATORY

Minimizing the loss function

$$L(\underline{w}) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 ; \quad \underline{w}^* = \underset{\underline{w}}{\operatorname{argmin}} L(\underline{w})$$

$$\text{Calculus} \Rightarrow \nabla_{\underline{w}} L(\underline{w}^*) = 0 . \Rightarrow \underline{w}^* = ?$$

observations (targets): $\underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, design matrix: $\underline{\underline{X}} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$ $(n \times 2)$ in this example

$$L = \| \underline{y} - \underline{\underline{X}} \cdot \underline{w} \|_2^2 = (\underline{y} - \underline{\underline{X}} \cdot \underline{w})^T (\underline{y} - \underline{\underline{X}} \cdot \underline{w})$$

bring the transpose inside

$$= (\underline{y}^T - \underline{w}^T \cdot \underline{\underline{X}}^T) \cdot (\underline{y} - \underline{\underline{X}} \cdot \underline{w})$$

$$= \underline{y}^T \cdot \underline{y} - \underline{y}^T \underline{\underline{X}} \cdot \underline{w} - \underline{w}^T \underline{\underline{X}}^T \underline{y} + \underline{w}^T \underline{\underline{X}}^T \cdot \underline{\underline{X}} \cdot \underline{w}$$

order matters

$$= \underline{y}^T \cdot \underline{y} - 2 \underline{w}^T \underline{\underline{X}}^T \underline{y} + \underline{w}^T \underline{\underline{X}}^T \underline{\underline{X}} \cdot \underline{w}$$

PREDICTIVE SCIENCE LABORATORY

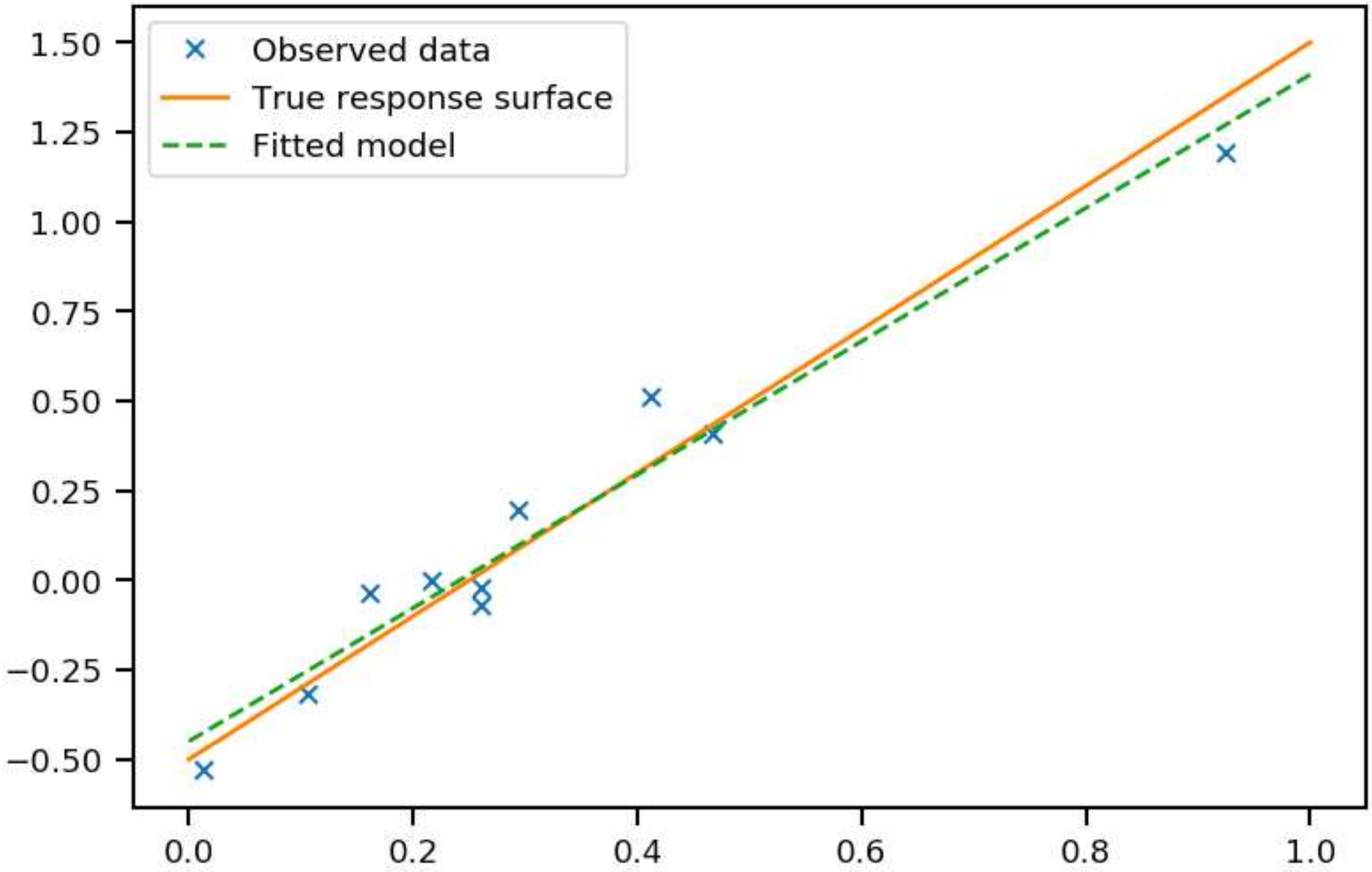
$$\nabla_{\underline{w}} L = -2 \underline{\underline{X}}^T \underline{y} + 2 \underline{\underline{X}}^T \underline{\underline{X}} \cdot \underline{w} = 0$$

$$\Rightarrow (\underline{\underline{X}}^T \cdot \underline{\underline{X}}) \underline{w} = \underline{\underline{X}}^T \underline{y}$$

$$\underline{\underline{A}} \cdot \underline{w} = \underline{b}$$

Solve the linear system

Example



PREDICTIVE
SCIENCE LABORATORY