

ENM 5310: Data-driven Modeling and Probabilistic Scientific Computing

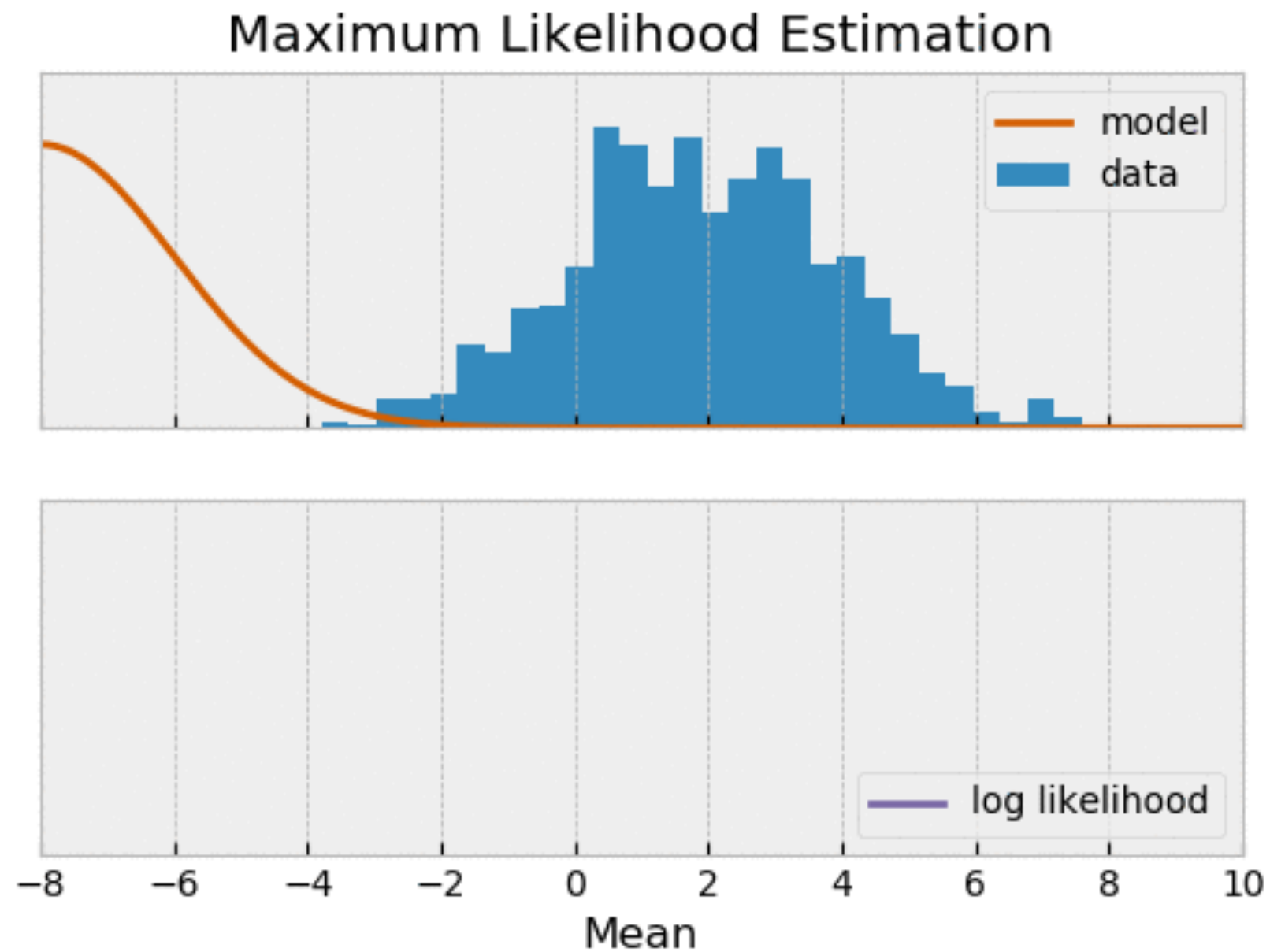
Lecture #4

Linear Regression



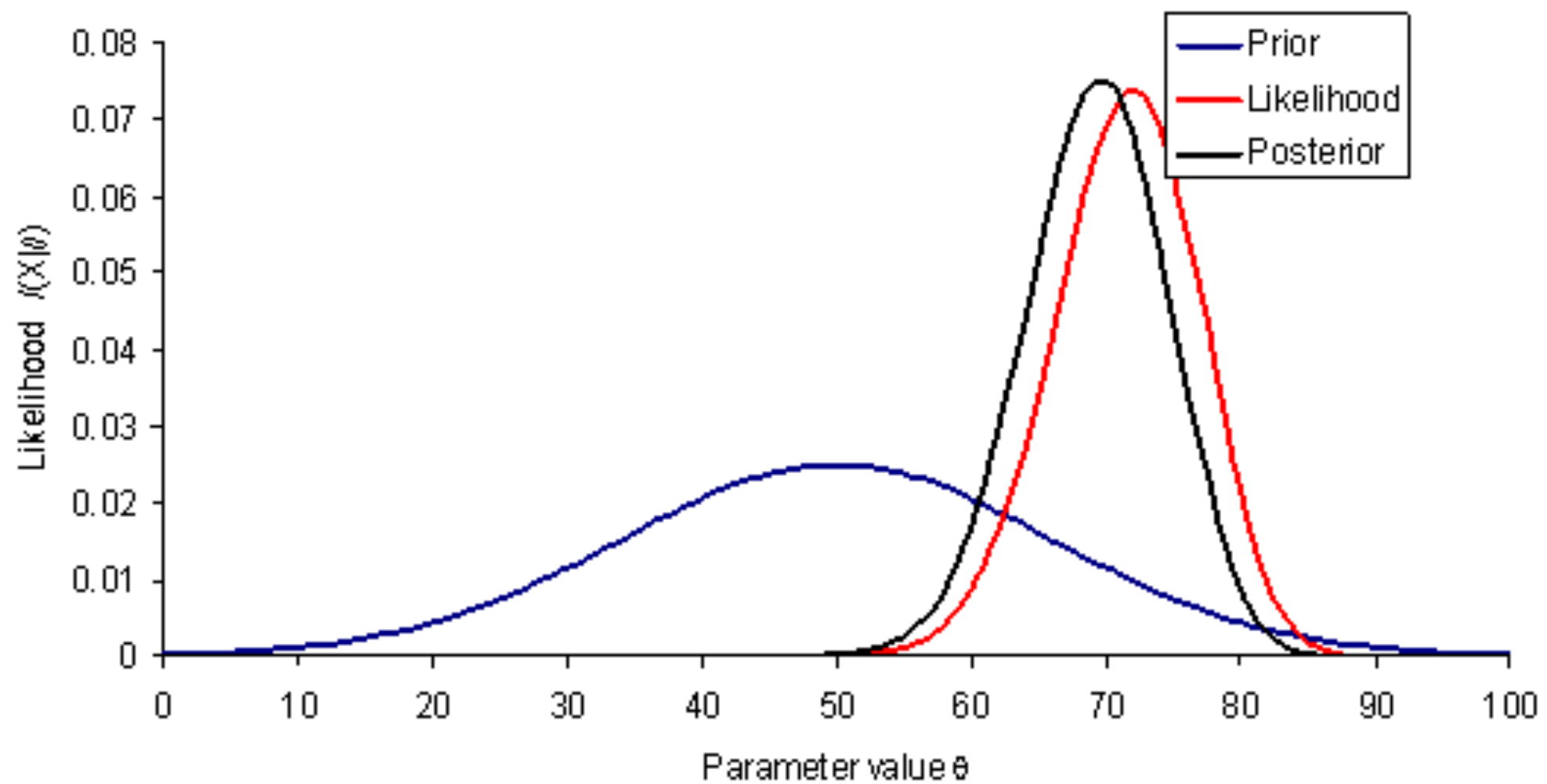
Maximum likelihood estimation

$$\theta_{\text{MLE}} = \arg \max_{\theta \in \Theta} p(\mathcal{D}|\theta)$$



Bayesian estimation

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})}$$



$$f : \mathcal{X} \rightarrow \mathcal{Y}$$

Supervised learning

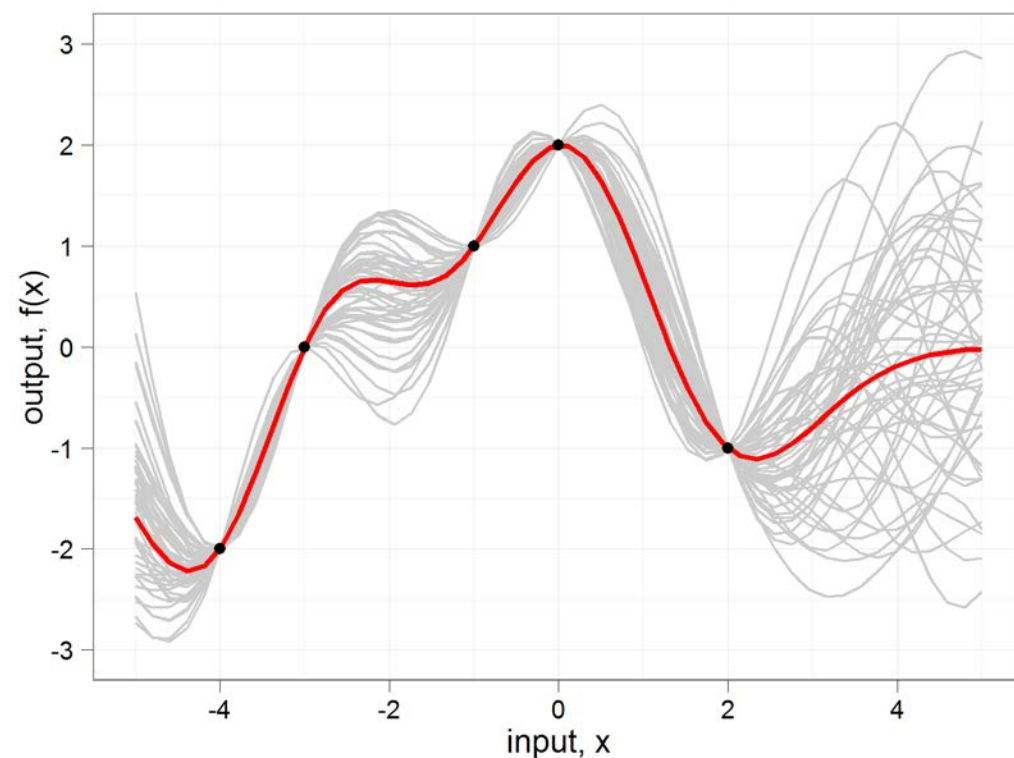
$$f : \mathcal{X} \rightarrow \mathcal{Y}$$

$$\mathcal{D} = \{x, y\}, \quad x \in \mathcal{X}, \quad y \in \mathcal{Y}$$

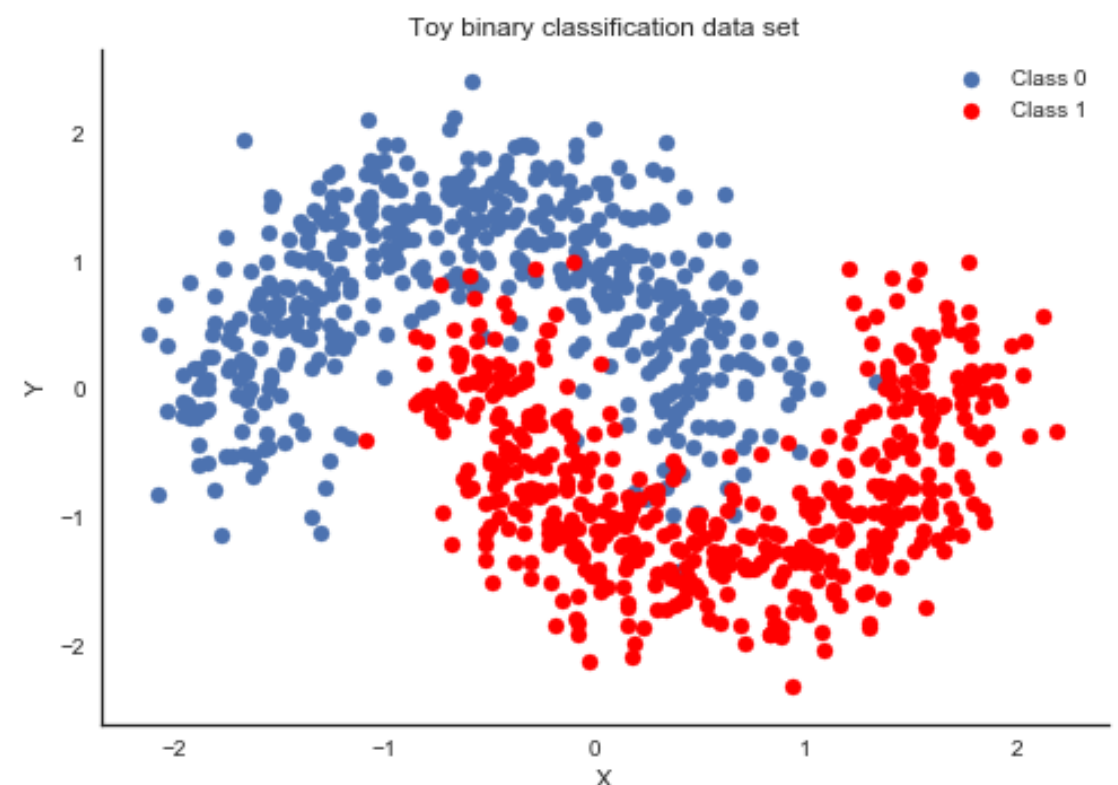
$$y = f(x) + \epsilon$$

$$p(f(x^*)|x^*, \mathcal{D})$$

Regression



Classification



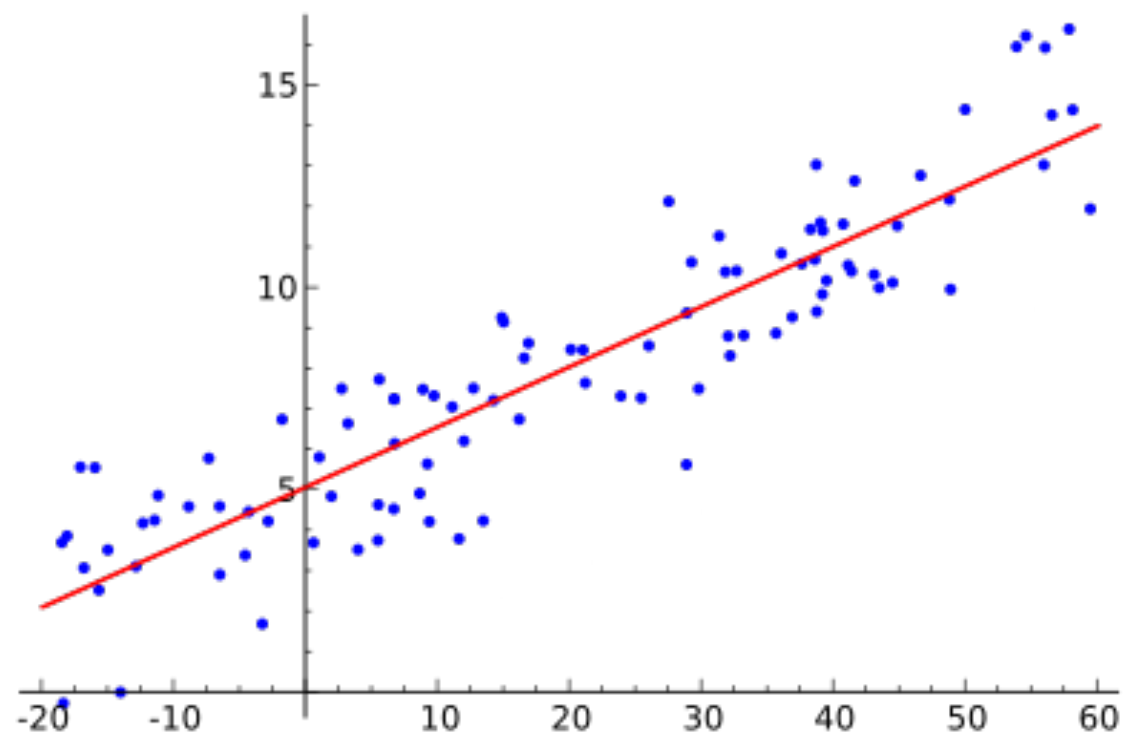
Linear regression

$$f : \mathcal{X} \rightarrow \mathcal{Y}$$

$$\mathcal{D} = \{x, y\}, \quad x \in \mathcal{X}, \quad y \in \mathcal{Y}$$

$$y = f(x) + \epsilon$$

$$f(x) = w^T x$$



“It’s not just about lines and planes!”

Linear regression with basis functions

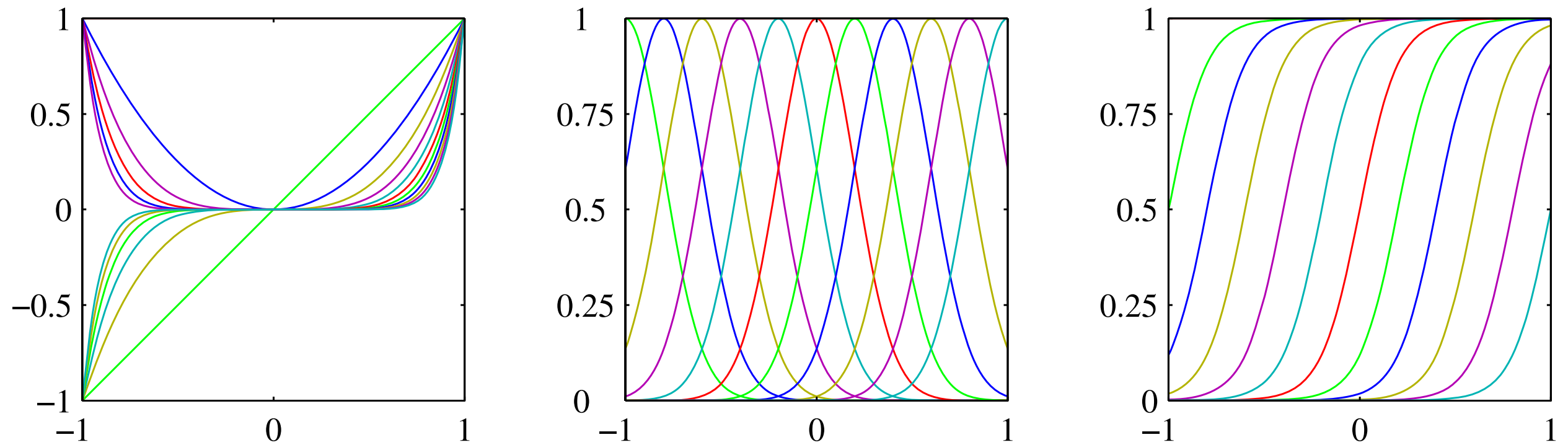
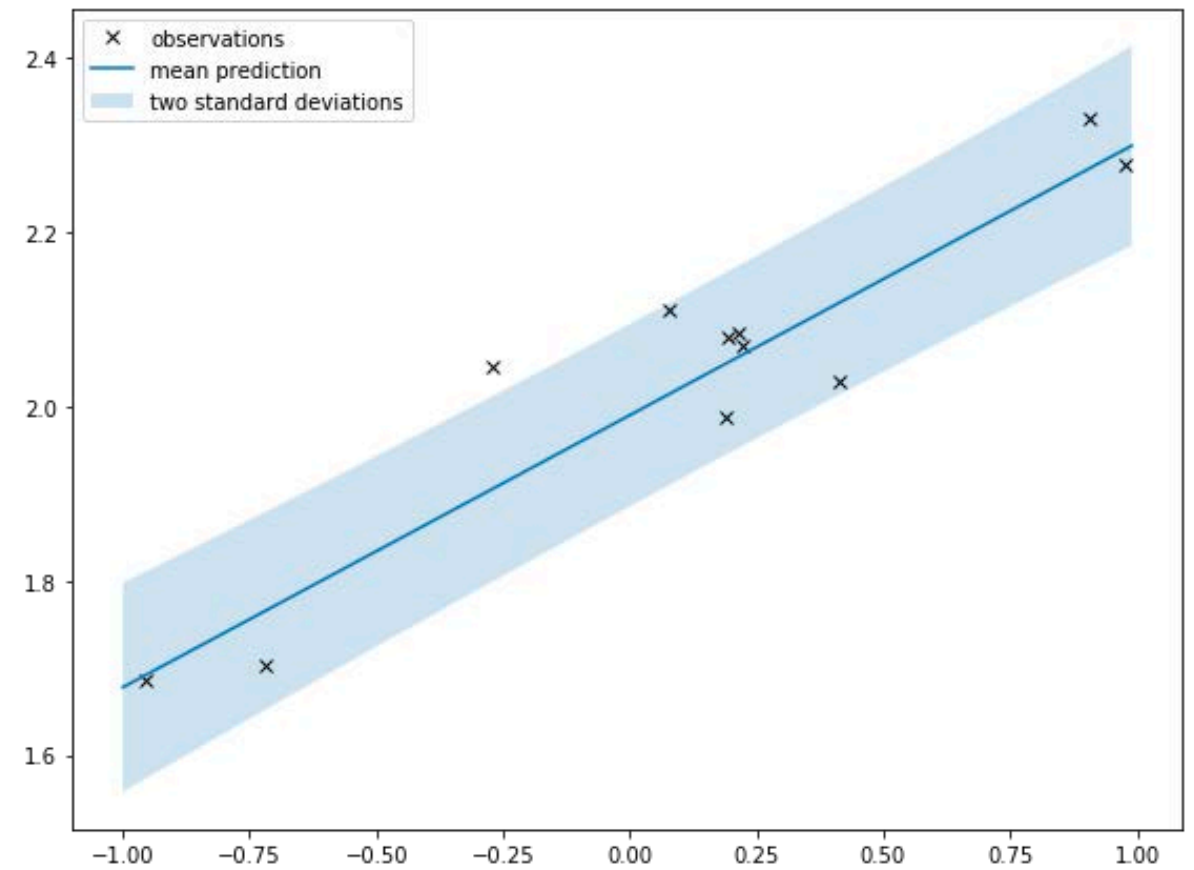
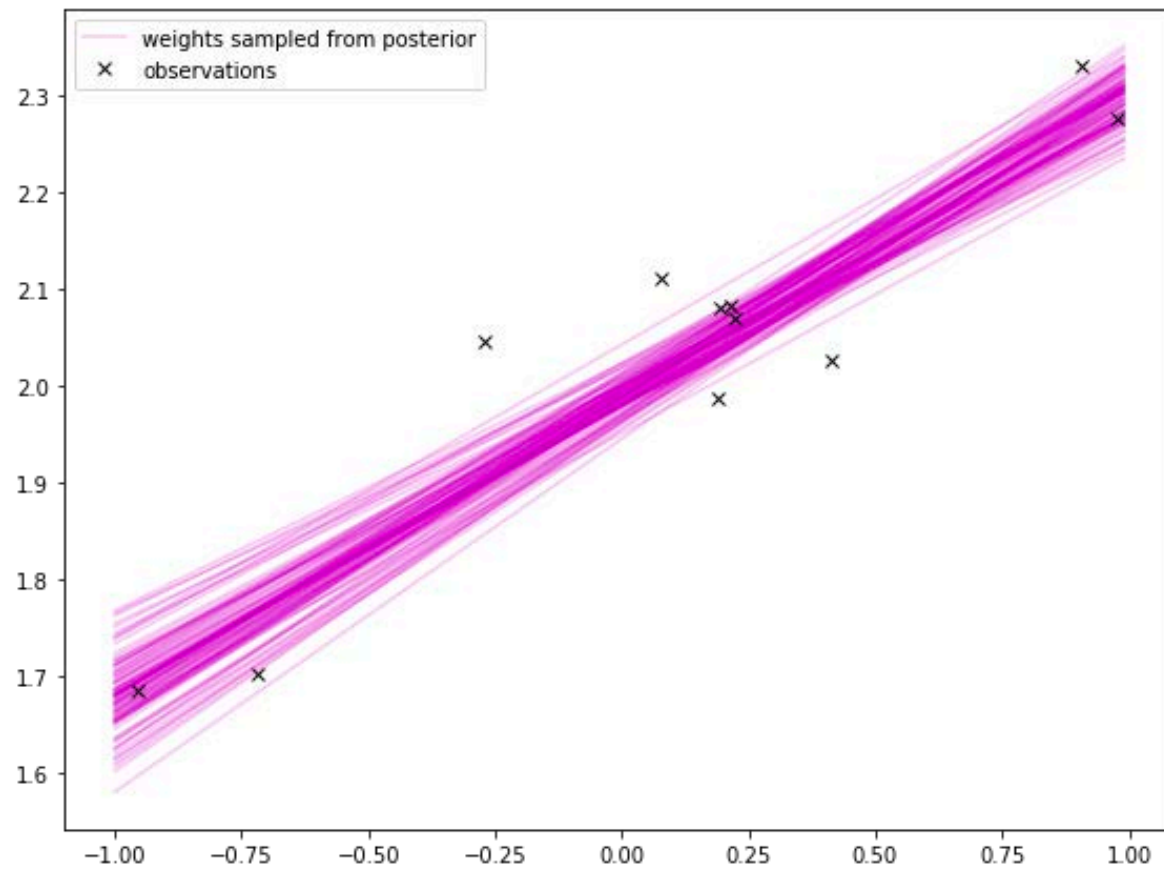
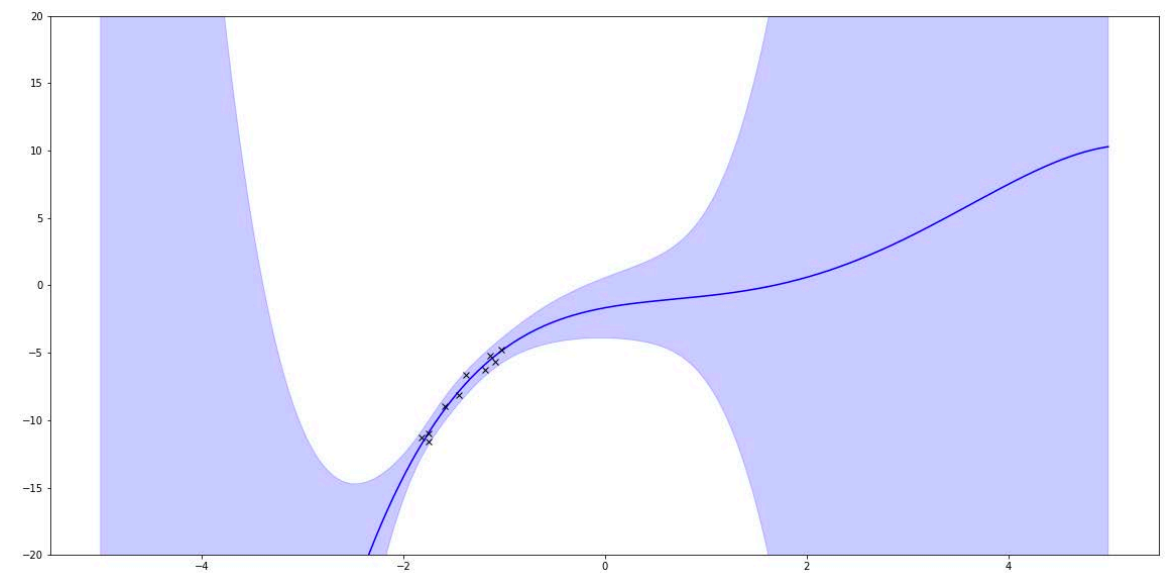
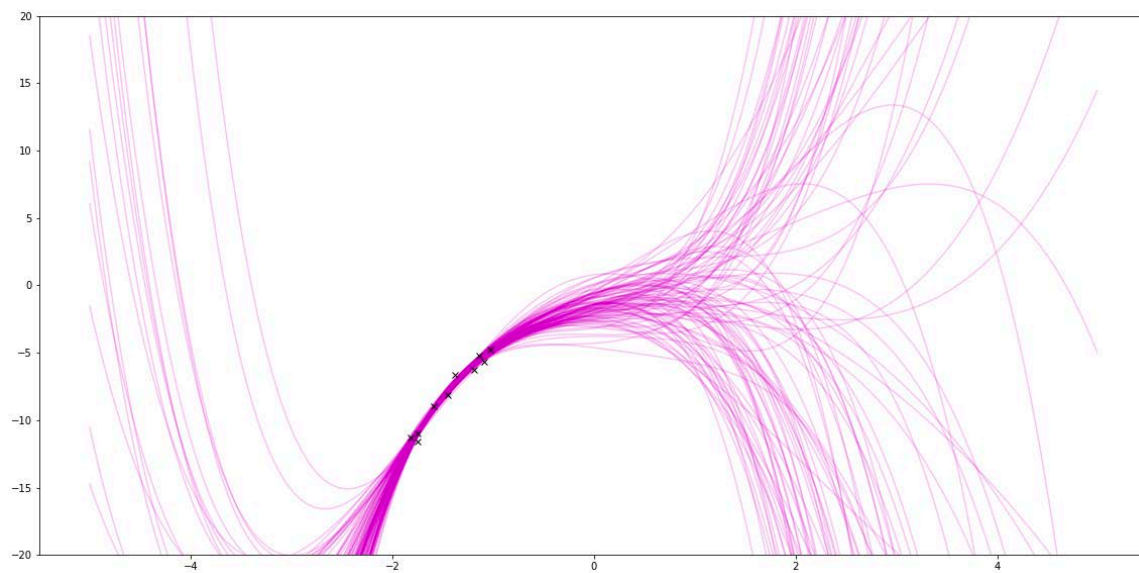


Figure 3.1 Examples of basis functions, showing polynomials on the left, Gaussians of the form (3.4) in the centre, and sigmoidal of the form (3.5) on the right.

Bayesian linear regression with basis functions



Nonlinear functions can be approximating using basis functions (or features)



$$\mathbf{y} = \mathbf{w}^T \phi(\mathbf{x}) + \epsilon$$

Geometrical interpretation

Figure 3.2 Geometrical interpretation of the least-squares solution, in an N -dimensional space whose axes are the values of t_1, \dots, t_N . The least-squares regression function is obtained by finding the orthogonal projection of the data vector \mathbf{t} onto the subspace spanned by the basis functions $\phi_j(\mathbf{x})$ in which each basis function is viewed as a vector φ_j of length N with elements $\phi_j(\mathbf{x}_n)$.

