

Lecture 14: Bayesian Linear Regression

Professor Ilias Bilionis

**Probabilistic interpretation of least
squares - Estimating the
measurement noise**

Reminder: Generalized linear model and least squares fit

$$\mathbf{x}_{1:n} = (x_1, \dots, x_n) ; \quad \mathbf{y}_{1:n} = (y_1, \dots, y_n) \quad (w_1, \dots, w_m)$$

$$y = w_1 \underbrace{\varphi_1(x)}_{\text{feature, basis function}} + \dots + w_m \underbrace{\varphi_m(x)}_{\text{feature, basis function}} = \sum_{j=1}^m w_j \varphi_j(x) = \underbrace{\varphi(x)^T}_{(\varphi_1(x), \dots, \varphi_m(x))} \underline{w}$$

$$\min L(\underline{w}) = \sum_{i=1}^n (y_i - \varphi(x_i)^T \underline{w})^2 \quad (y_1, \dots, y_n)$$

$$\nabla_{\underline{w}} L(\underline{w}) = 0 \Rightarrow \underbrace{\Phi^T \Phi}_{\text{Design matrix } n \times m} \cdot \underline{w} = \underbrace{\Phi^T}_{\text{Design matrix } n \times m} \cdot \underline{y}$$

$$\Phi = \begin{pmatrix} \varphi_1(x_1) & \dots & \varphi_m(x_1) \\ \vdots & & \vdots \\ \varphi_1(x_n) & & \varphi_m(x_n) \end{pmatrix}$$

