Benign Overfitting

Linus Böhm, Jurek Rostalsky

1 Basic definitions

Let $x \in \mathbb{H}, y \in \mathbb{R}$ with zero mean. Where $\mathbb{H}$ is a Hilbert space.

Definition 1.1 (covariance matrix)**.**

$$\Sigma = \mathbb{E}\left((x - \mathbb{E}(x))\,(x - \mathbb{E}(x))^T\right) = \mathbb{E}(xx^T) \tag{1}$$

Definition 1.2 (linear regression)**.** The problem of finding a parameter vector $\theta^* \in \mathbb{H}$ with

$$\theta^* = arg\min_{\theta} \mathbb{E}\left((y - x^T\theta)^2\right) \tag{2}$$

is called **linear regression**.

Let $((x_1, y_1), ..., (x_n, y_n)) \in (\mathbb{H} \times \mathbb{R})^n$ a list of n sampled data points. Now we define the matrix $X = \left(x_1\, x_2\, ...\, x_n\right)$ and the vector $y = (y_1\, y_2\, ...\, y_n)^T$. If there is a $\theta \in \mathbb{H}$ with $y - X^T\theta = 0$ that θ is a minimum of the linear regression problem sind the expectation of a square is non negative. Usually such a θ isn't unique, so we are interested in the minimum norm θ with that property.

Definition 1.3 (minimum norm estimator)**.** For given samples $X \in \mathbb{H}^n, y \in \mathbb{R}^n$. The **minimum norm estimator** θ is the solution of the QQP:

$$\hat{\theta} = arg\min_{\theta} \|\theta\|^2 \qquad \text{subject to: } \left\|X^T\theta - y\right\|^2 = \min_{\beta} \left\|X^T\beta - y\right\|^2 \tag{QQP}$$

The minimum norm estimator can be obtained by solving the normal equation:

$$XX^T\theta = Xy, \tag{3}$$

which can be done by numerical stable with QR-decomposition.