

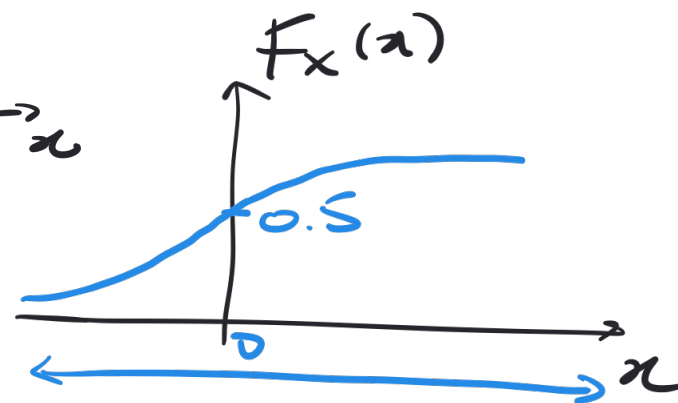
Q-Q (Quantile - Quantile) Plot

Continuous R.V. X : PDF $f_X(x)$



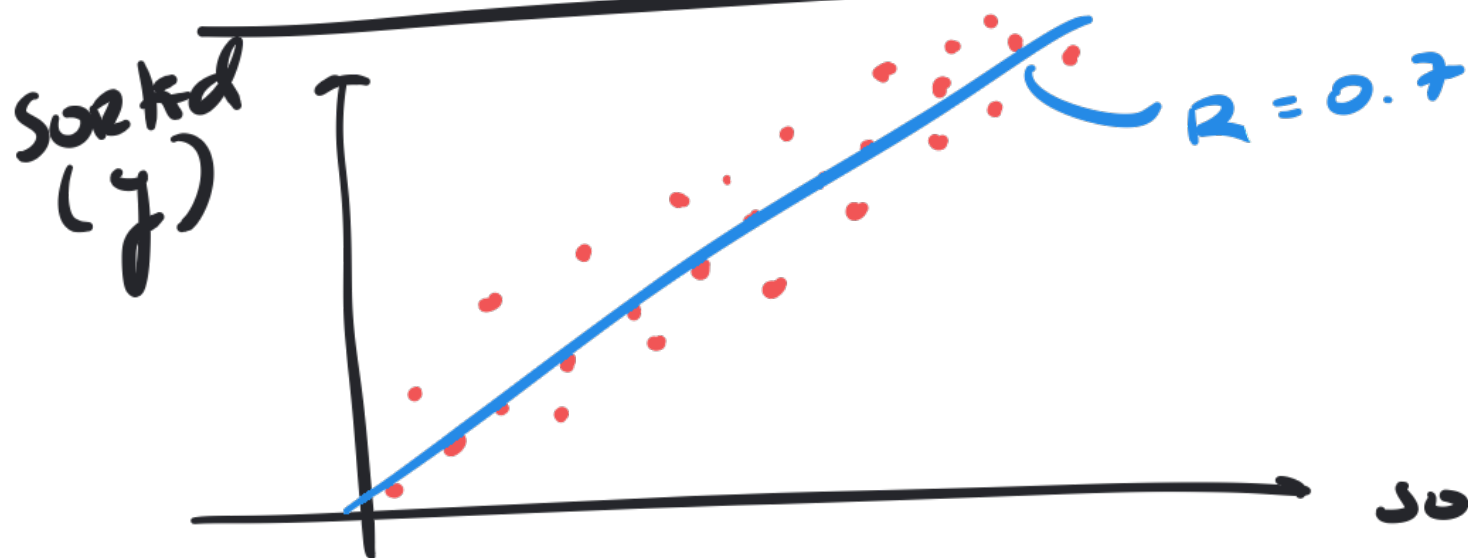
CDF : $F_X(x) = P(X \leq x)$

$F_X(x=0) = 0.5$ if X is Gaussian



target $t \leftarrow$ drawn from R.V.

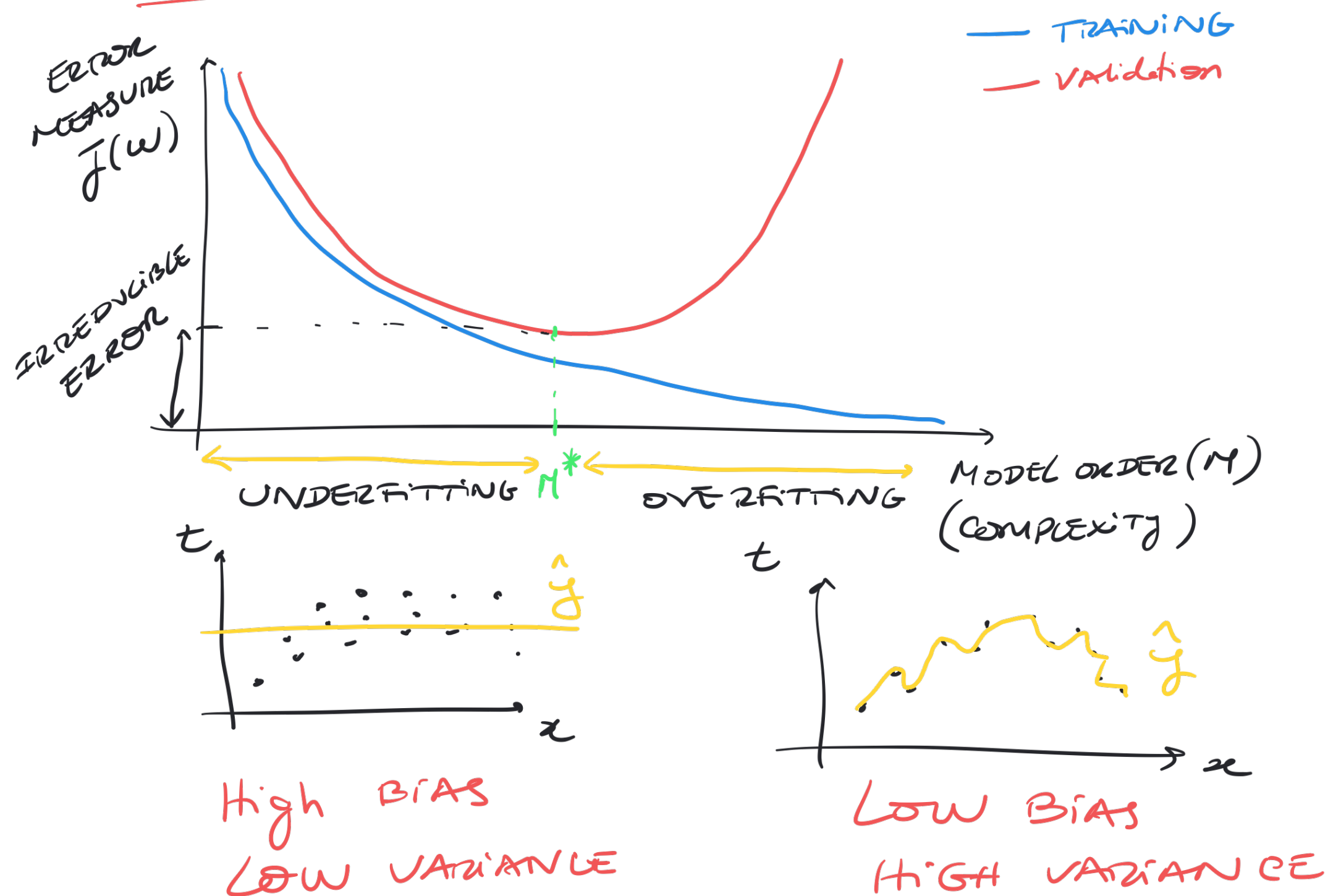
predictions $y \leftarrow$ should (ideally) approx.
the same prob. mode



Coefficient of
Determination
 $R^2 = 0.7^2$

In practice $R^2 \gg 0.9$ to be considered
successful.

The Bias-Variance Trade-Off



$y \equiv \text{true (unknown model)}$

$\hat{y} \equiv \text{Estimator}$

OBJ. FCT:

$$J(w) = \frac{1}{2} \sum_{i=1}^N (t_i - \hat{y}_i)^2 = \frac{1}{2} (t - y)^2$$

EXPECTED VALUE ^{$\equiv \text{MEAN}$} of $J(w)$ BASED ON
DATASET AND Hyper parameter choices

$$E[J(w)] = E[(t - y)^2] = (*)$$

$$\textcircled{*} = E[(t - y)^2]$$

$$= E[(t - E_D[y] + E_D[y] - y)^2]$$

$E_D[y] \equiv$ true
Mapper on
DATASET D

$$= \underbrace{E[(t - E_D[y])^2]}_{\text{VARIANCE of ESTIMATOR}} + \underbrace{E[(E_D[y] - y)^2]}_{\text{BIAS}^2 \text{ of ESTIMATOR}} + \underbrace{E[2(t - E_D[y])(E_D[y] - y)]}_{\text{IRREDUCIBLE ERROR}}$$

Key:

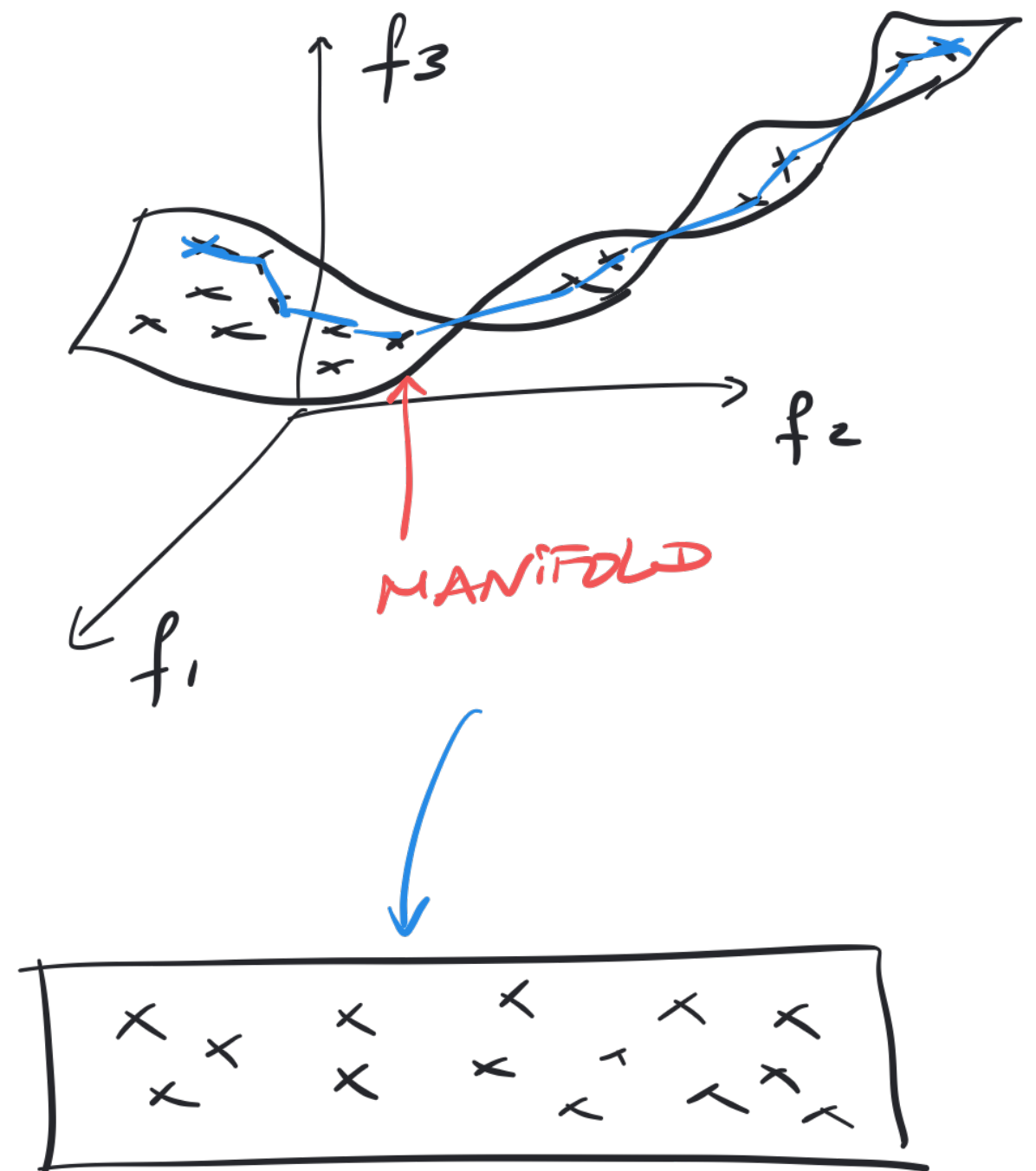
t — target

$E_D[y] = \hat{y}$ — ESTIMATOR

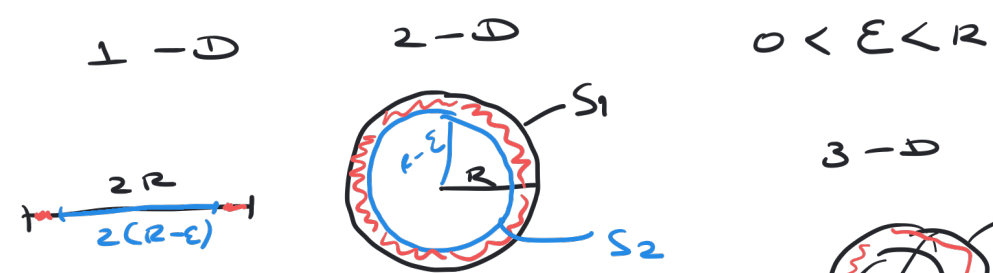
y — TRUE MAPPER

The Curse of Dimensionality

Things do not
behave as we
expect them to
in high-
dimension!!



Example: volume of crust



volume of a sphere
of radius R
in D -dimensions

$$\frac{R^D \cdot \pi^{D/2}}{\Gamma(D/2 + 1)} \quad (\Gamma(n) = (n-1)!)$$

$$\begin{aligned} \text{Ratio} &= \frac{V_{\text{CRUST}}}{V_{S1}} = \frac{V_{S1} - V_{S2}}{V_{S1}} = 1 - \frac{V_{S2}}{V_{S1}} \\ &= 1 - \frac{\frac{(R-\epsilon)^D \cdot \pi^{D/2}}{\Gamma(D/2 + 1)}}{\frac{R^D \cdot \pi^{D/2}}{\Gamma(D/2 + 1)}} = 1 - \frac{(R-\epsilon)^D}{R^D} \\ &= 1 - \frac{R^D \left(1 - \frac{\epsilon}{R}\right)^D}{R^D} = 1 - \left(1 - \frac{\epsilon}{R}\right)^D \xrightarrow{D \rightarrow \infty} 1 \end{aligned}$$

APPROACH TO TAKE :

① Dimensionality Reduction:

A) FEATURE SELECTION

E.g. RECURSIVE FEATURE
FORWARD FEATURE

ELIMINATION
SELECTION

B) FEATURE EXTRACTION

E.g. PCA
LDA

② MANIFOLD LEARNING