Helwan University
Faculty of Computers and Information

(Computer Exercise)

Department of Computer Science
Pattern Recognition

A First Simulation Example on Designing and Assessing a Regression Function (cont.)

- Similar to the previous simulation on linear model, this time assume that Y is related to X as follows

$$Y = \sin X + \varepsilon,$$

where X is uniformally distributed in $[-\pi, \pi]$; and $\varepsilon \sim \mathcal{N}(0, 0.1)$. Notice that, in this case $\sigma^2_{Y|X}$ is constant.

- Build different models, intercept, first order, second order, ..., p^{th} order.
- For each model generate the following figures:
  - a figure showing the true model, $\mathrm{E}\, Y|X = \sin X$, along with the 500 fits.
  - a figure showing (vs. n_{tr}): the risk of the true model $\mathrm{E}_{x_0}\, \sigma^2_{Y|X=x_0}$, the average bias $\mathrm{E}_{x_0}\, Bias^2(\widehat{y}_0, y_0)$, the average variance $\mathrm{E}_{x_0}\, \mathrm{Var}_{\mathbf{tr}}(\widehat{y}_0)$, and the mean error $\mathrm{E}_{x_0}\, \mathrm{E}_{\mathbf{tr}}\, err_{\mathbf{tr}}(x_0) = \mathrm{E}_{\mathbf{tr}}\, \mathrm{E}_{x_0}\, err_{\mathbf{tr}}(x_0) = \mathrm{E}_{\mathbf{tr}}\, err_{\mathbf{tr}}$; verify that

$$\underset{\mathbf{tr}}{\mathrm{E}}\, err_{\mathbf{tr}} = \underset{x_0}{\mathrm{E}}\, \sigma^2_{Y|X=x_0} + \underset{x_0}{\mathrm{E}}\, Bias^2(\widehat{y}_0, y_0) + \underset{x_0}{\mathrm{E}}\, \underset{\mathbf{tr}}{\mathrm{Var}}(\widehat{y}_0).$$

- For each $n_{\mathbf{tr}}$, generate a figure (vs. the complexity p), that shows the L.H.S and the three terms of the R.H.S. of the equation above.
- For each $n_{\mathbf{tr}}$, what is the best model for this problem?