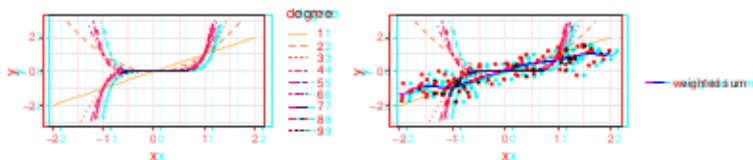


POLYNOMIAL REGRESSION

- Simple & flexible choice for basis funs: ***d*-polynomials**
- Idea: map x_j to (weighted) sum of its monomials up to order $d \in \mathbb{N}$

$$\phi^{(d)} : \mathbb{R} \rightarrow \mathbb{R}, \quad x_j \mapsto \sum_{k=1}^d \beta_k x_j^k$$



- How to estimate coefficients β_k ?
 - Both LM & polynomials **linear** in their params \rightsquigarrow merge
 - E.g., $f(\mathbf{x}) = \theta_0 + \theta_1 \phi^{(d)}(x) = \theta_0 + \sum_{k=1}^d \theta_{1,k} x^k$

$$\rightsquigarrow \mathbf{X} = \begin{pmatrix} 1 & x^{(1)} & (x^{(1)})^2 & \dots & (x^{(1)})^d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x^{(n)} & (x^{(n)})^2 & \dots & (x^{(n)})^d \end{pmatrix}, \quad \boldsymbol{\theta} \in \mathbb{R}^{d+1}$$

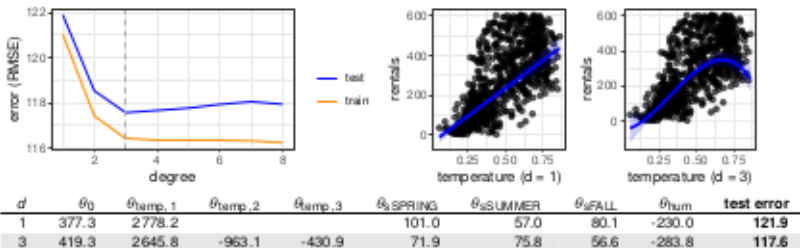


BIKE RENTAL EXAMPLE

- OpenML task `dailybike`: predict rentals from weather conditions
- Hunch: non-linear effect of temperature \rightsquigarrow include with polynomial:

$$f(\mathbf{x}) = \sum_{k=1}^d \theta_{\text{temperature},k} X_{\text{temperature}}^k + \theta_{\text{season}} X_{\text{season}} + \theta_{\text{humidity}} X_{\text{humidity}}$$

- Test error² confirms suspicion \rightsquigarrow minimal for $d = 3$



- Conclusion: flexible effects can improve fit/performance

²Reliable insights about model performance only via separate test dataset not used during training (here computed via 10-fold *cross validation*). Much more on this in Evaluation chapter.

