

Q6 Explain the general model of linear regression with different possible regularizations.

- $g(x) = \sum_{i=0}^{M-1} \phi_i(x) w_i$  general formula for linear regression.  $\phi \rightarrow$  Base functions,  $w \rightarrow$  weights
- Choose basis functions such that data is  $\approx$  linear in feature space
- Linear regression typically overfits. Low bias, high variance (overfitting)
- Tradeoff you add regularization to get lower variance but higher bias.
- Methods for regularization: Ridge (L2) and Lasso (L1).
- L2:  $R(w) = \frac{1}{2} \|w\|_2^2$ , L1:  $R(w) = \|w\|_1$ , for objective function:  $J(w) = \frac{1}{2} \|y - \phi(w)\|^2 + \lambda R(w)$   
 $\lambda$  controls the regularization strength. Add penalty term.
- L2: Closed form solution, Gaussian prior distribution, no feature selection
- L1: Iterative optimization, Laplace - "yes" - , yes - "no" -

