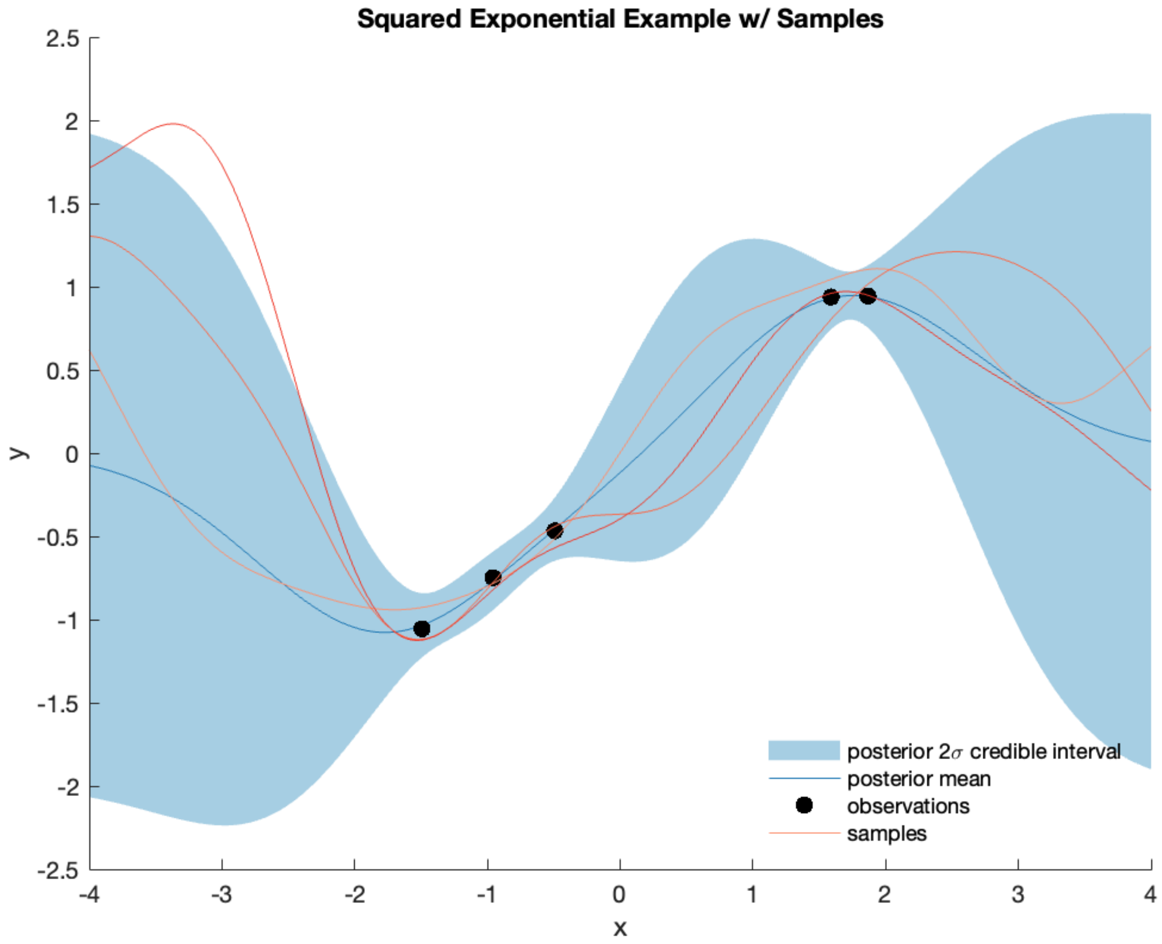
10-424/624: Bayesian Methods in ML Lecture 12: Supplement & Figures

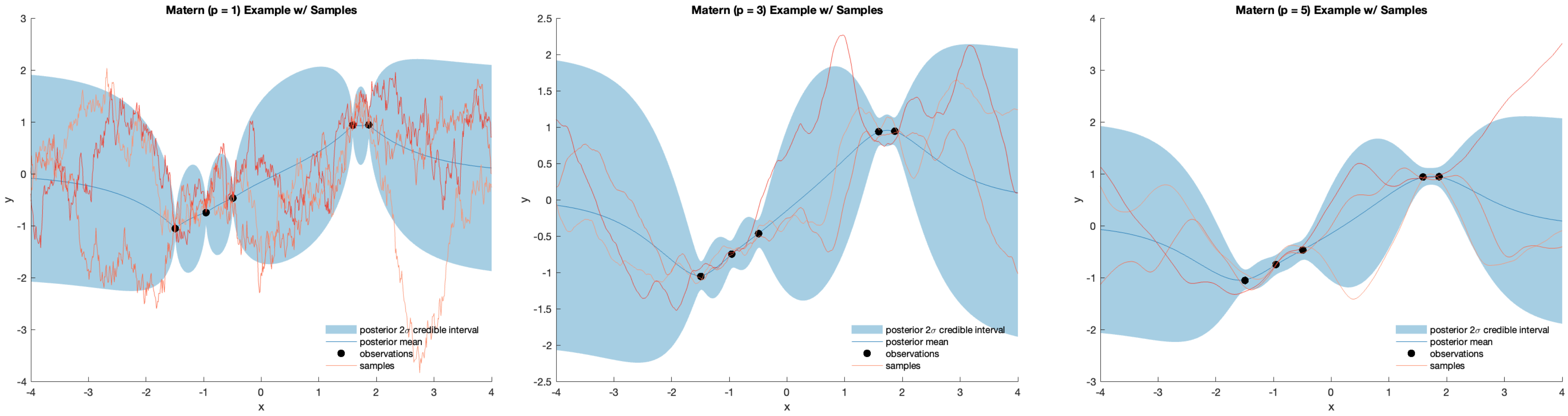
Henry Chai

2/20/25

Squared Exponential Kernel

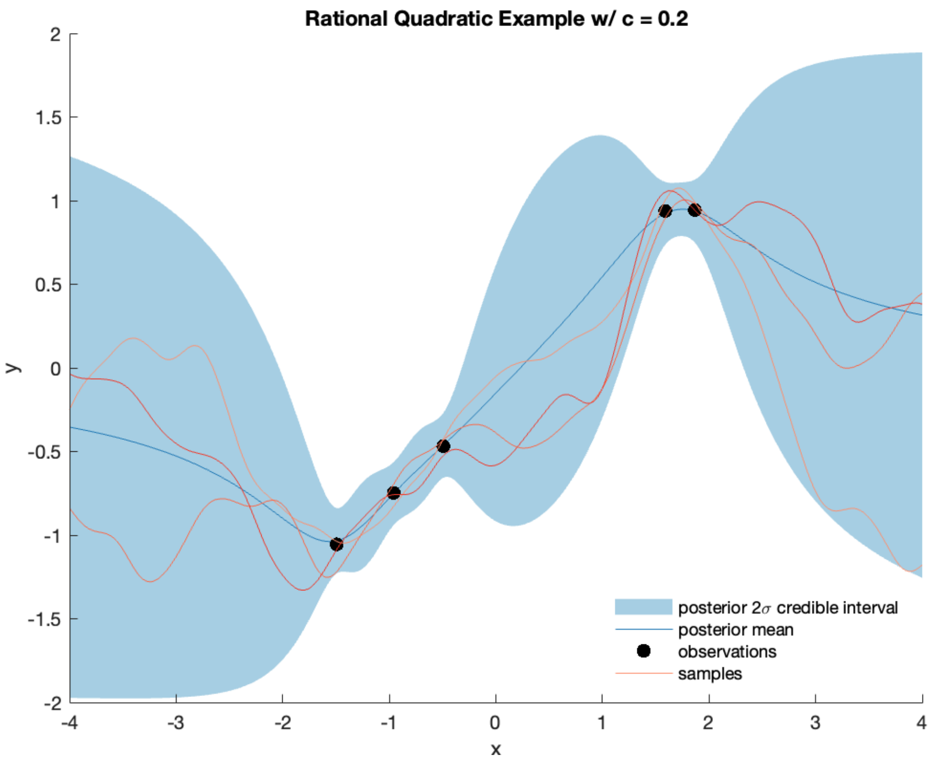


A Gaussian process posterior with the squared exponential kernel fit to 5 observations; the kernel uses $\ell = \lambda = 1$ and $\sigma = 0.1$. 3 sample paths from the posterior are also shown

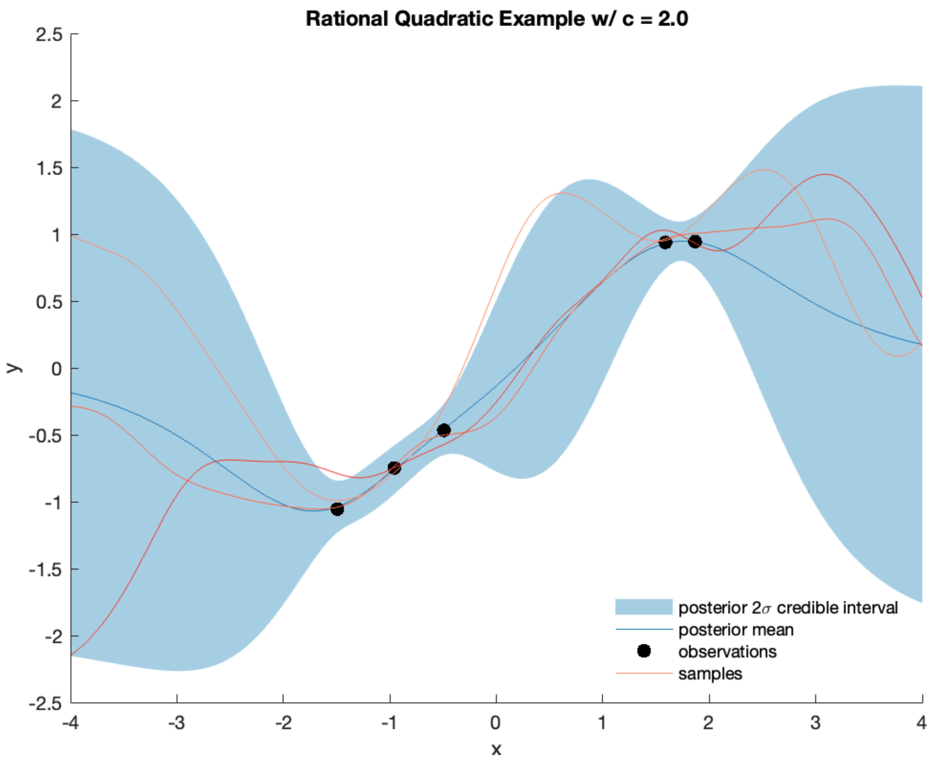


$\nu = {}^{1}\!/_{2}$ $\nu = {}^{3}\!/_{2}$ $\nu = {}^{5}\!/_{2}$

Matern Covariance Functions



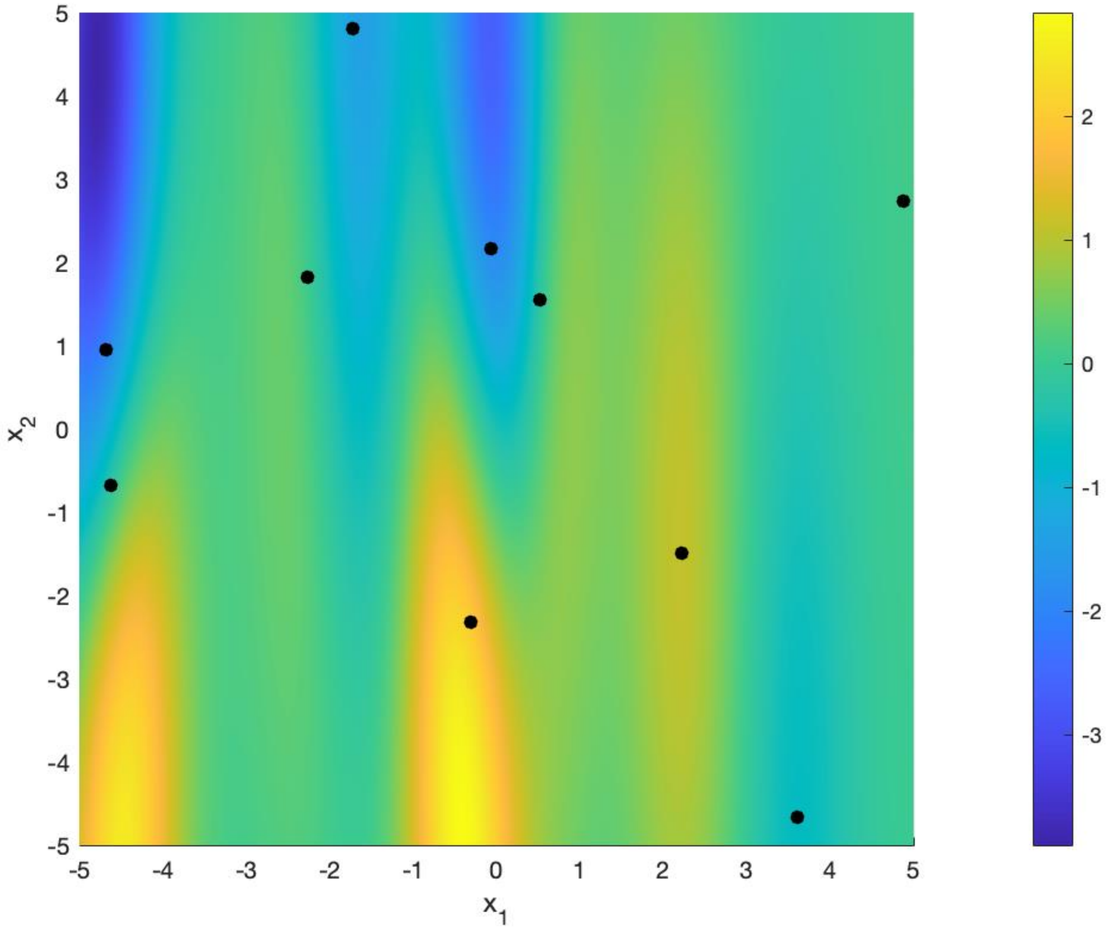
$c = 0.2$



$c = 2$

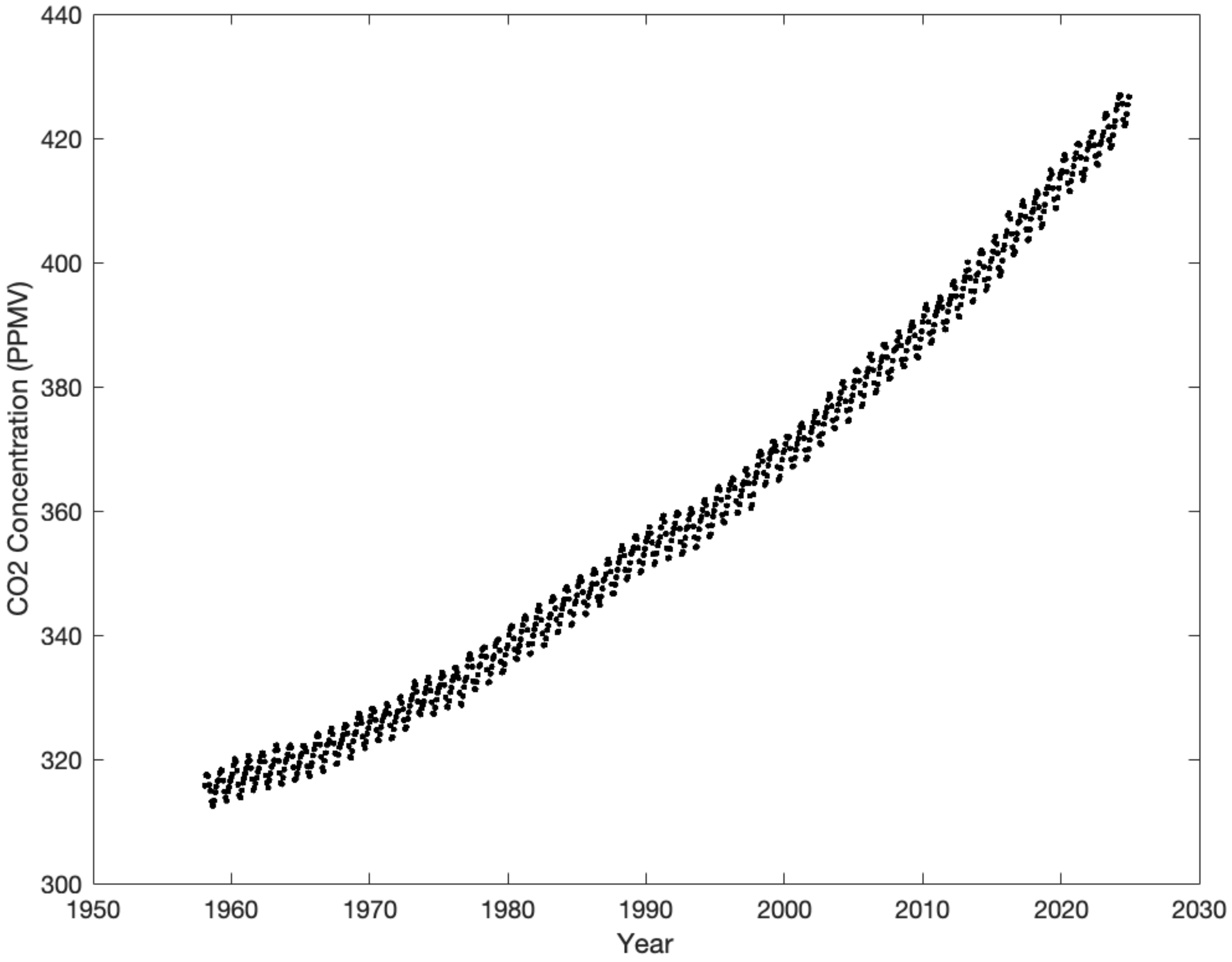
Rational Quadratic Covariance Function

Automatic Relevance Determination



A Gaussian process posterior mean with the squared exponential kernel using automatic relevance determination where $\ell_1 = 0.5$ and $\ell_2 = 5$; the kernel uses $\lambda = 1$ and $\sigma = 0.1$.

Mauna Loa CO2 Concentrations



Combining Kernels

- a squared exponential kernel: $k_1(x, x') = \theta_1^2 \exp\left(-\frac{(x-x')^2}{2\theta_2^2}\right)$,
- a squared exponential kernel multiplied by a *periodic* kernel:
  $k_2(x, x') = \theta_3^2 \exp\left(-\frac{(x-x')^2}{2\theta_4^2} - \frac{2\sin(\pi(x-x'))}{\theta_5^2}\right)$,
- a rational quadratic kernel: $k_3(x, x') = \theta_6^2 \left(1 + \frac{(x-x')^2}{2\theta_8\theta_7^2}\right)^{\theta_8}$, and
- another squared exponential kernel: $k_4(x, x') = \theta_9^2 \exp\left(-\frac{(x-x')^2}{2\theta_{10}^2}\right)$.

The final kernel is simply $k(x, x') = k_1(x, x') + k_2(x, x') + k_3(x, x') + k_4(x, x')$; the hyperparameters $\{\theta_1, \theta_2, \ldots, \theta_{10}\}$ are fit to maximize the likelihood of the data between 1958 and 2003.

Combining Kernels