Bayesian Model Selection: Prior vs Posterior over Polynomial Degree

Overview

This document explains the meaning of the plot `figures/prior_vs_posterior.png`, which visualizes the effect of Bayesian model selection using a prior over polynomial degrees.

Purpose of the Plot

The goal is to evaluate how a Bayesian framework updates our belief about which model degree best explains the data. Given a prior belief $P(m)$ over polynomial degrees m, and the evidence from test set chi-squared values $\chi^2_B(m)$, we compute the posterior distribution:

$$P(m \mid D) \propto \exp\left(-\frac{1}{2\sigma^2}\chi^2_B(m)\right) \cdot P(m)$$

Here:

- $P(m)$ is the exponential prior: $P(m) \propto e^{-\lambda m}$
- $\chi^2_B(m)$ is the cross-validated chi-squared on dataset D_B
- σ^2 is the known variance of the noise

Interpretation

The plot compares:

- The prior distribution (black dashed line), reflecting our initial skepticism toward complex models
- The posterior distribution (blue solid line), showing which degrees are most probable *after seeing the data*

A peak in the posterior identifies the most likely polynomial degree under the Bayesian model. If the posterior is sharply peaked, it indicates confident selection of a specific model; a flatter posterior reflects model uncertainty.

Conclusion

This plot illustrates the integration of prior beliefs and data evidence in model selection. It highlights how increasing model complexity is penalized unless strongly supported by improved fit (lower χ^2).