Summary of Chapter 7: Bayesian Linear Regression

Chapter 7 introduces Bayesian linear regression, a probabilistic approach to linear regression that provides uncertainty estimates and regularization.

1. **Introduction to Bayesian Linear Regression**:
   - Bayesian linear regression models the relationship between input variables and output variables using a linear function, while incorporating prior knowledge and uncertainty about model parameters.
   - It provides a probabilistic framework for estimating model parameters and making predictions, along with uncertainty estimates.
2. **Bayesian Inference**:
   - In Bayesian inference, we compute the posterior distribution over model parameters given observed data using Bayes' theorem:

     $$p(\theta|\mathbf{X}, \mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{X}, \theta)p(\theta)}{p(\mathbf{y}|\mathbf{X})}$$

     where θ represents the model parameters, $\mathbf{X}$ is the input data, and $\mathbf{y}$ is the output data.
3. **Posterior Predictive Distribution**:
   - The posterior predictive distribution provides a distribution over predictions given observed data and uncertainty in model parameters:

     $$p(\mathbf{y}^*|\mathbf{X}^*, \mathbf{X}, \mathbf{y}) = \int p(\mathbf{y}^*|\mathbf{X}^*, \theta)p(\theta|\mathbf{X}, \mathbf{y})\, d\theta$$

     where $\mathbf{X}^*$ represents new input data and $\mathbf{y}^*$ represents corresponding predictions.
4. **Bayesian Linear Regression Models**:
   - Different priors can be used to regularize the model and encode prior knowledge about model parameters.
   - Common choices include Gaussian priors for weights and hyperparameters, resulting in a Gaussian posterior distribution.
   - Extensions include hierarchical models, where hyperparameters are also treated as random variables with their own priors.
5. **Bayesian Model Selection**:
   - Bayesian model selection involves comparing different models by computing their marginal likelihoods or evidence.

- The marginal likelihood represents the probability of the observed data under a particular model, integrating over all possible values of model parameters:

$$p(\mathbf{y}|\mathbf{X}, \mathcal{M}) = \int p(\mathbf{y}|\mathbf{X}, \theta, \mathcal{M}) p(\theta|\mathcal{M}) \, d\theta$$

  where $\mathcal{M}$ represents the model.

6. **Extensions and Applications**:

- Bayesian linear regression can be extended to handle nonlinear relationships using basis functions or kernel methods.
- It finds applications in various fields, including economics, finance, engineering, and machine learning, where uncertainty estimates and regularization are important.