Computational Statistics
Gibbs sampling: application to linear regression

The file `investment.txt` contains 15 yearly observations of U.S. investment data for the period 1968-1982. The variables are

- `Year` = Date,
- `GNP` = Nominal GNP,
- `Invest` = Nominal Investment,
- `CPI` = Consumer price index,
- `Interest` = Interest rate,
- `Inflation` = rate of inflation computed as the percentage change in the CPI.

1. We consider the linear regression model $\mathbf{y} \sim \mathcal{N}_n(X\boldsymbol{\beta}, \sigma^2 I_n)$ with the dependent variable `Invest/CPI` and, as covariates, the time trend (a vector of integers from 1 to 15), `GNP/CPI`, `Interest` and `Inflation`. Plot the data and compute the least-squares estimate of the parameters.

2. Program a Gibbs sampler to simulate the posterior distribution of the parameters $\boldsymbol{\beta}$ and σ, assuming a normal prior for $\boldsymbol{\beta}$ and an inverse Gamma prior for σ. To simulate from the multivariate normal distribution, you can use the function `mvrnorm` in the package `MASS`.

3. Use the Gibbs sampler with different prior probability distributions. To determine the parameters of the inverse Gamma distribution, you may use the following equations that relate the shape parameter α and the rate parameter δ to the mean μ and the variance V of the inverse Gamma distribution:

$$\alpha = \mu^2/V + 2, \quad \delta = \mu(\mu^2/V + 1).$$

4. Write a function that generates samples from the posterior predictive distribution of $y_0 = \mathbf{x}_0^T\boldsymbol{\beta} + \sigma u$, with $u \sim \mathcal{N}(0, 1)$ and from the posterior distribution of $E(y_0) = \mathbf{x}_0^T\boldsymbol{\beta}$.