Assignment 6.

Given December 10, due December 31.

Objective: To explore numerical optimization and statistical estimation.

A very simple ARCH/GARCH type model is

$$\sigma_{k+1}^2 = \alpha \sigma_k^2 + \beta Z_k^2 \ , \tag{1}$$

$$X_{k+1} = X_k + \gamma \sigma_k Z_k \ , \tag{2}$$

where the Z_k are independent standard normal ransom variables. This is a different approach to stochastic volatility that allows for bursts of high volatility.

First we want to construct maximum likelihood estimates of the parameters α, β, and γ. Assume that $X_0 = 0$ and $\sigma_0 = 1$. For given values of the parameters, the probability density for a specific sequence $\vec{x} = (x_1, \ldots, x_n)$ is

$$f(\vec{x}, \alpha, \beta, \gamma) = \frac{1}{(2\pi)^{n/2} \gamma^n \sigma_0 \sigma_1 \cdots \sigma_{n-1}} \exp\left(- \sum_{k=0}^{n-1} \frac{(x_{k+1} - x_k)^2}{2\gamma^2 \sigma_k^2} \right) \tag{3}$$

The numbers σ_k in (3) can be computed, given $\vec{x}$, from (1) and $\sigma_0 = 1$. This is how f comes to depend on the parameter β. The maximum likelihood estimates $\hat{\alpha}(\vec{x})$, $\hat{\beta}(\vec{x})$, and $\hat{\gamma}(\vec{x})$, are found by maximizing f over α, β, and γ.

The simplest way to do this three dimensional optimization problem is to choose search directions using the gradient method and then use bisection search. This requires you to compute first derivatives of f with respect to the parameters. Make sure you program is modular: you should test the line search part separately, for example. It will probably work better if you optimize $\log(f)$ instead, the answer will be the same but the objective function will be less wild. Are there local maxima that are not the global maximum?

Second, we will make ad-hoc error bars for the estimated parameter values by a version of the jackknife process. Produce artificial data $\vec{x}_1$, …, $\vec{x}_M$ by running the model (1), (2) using the estimated parameter values. Then use the optimization program on these artificial data sets to compute $\hat{\alpha}(\vec{x}_l)$ for $l = 1$, …, l_M. The interval that includes 95% of these artificial α estimates is a poor man's 95% confidence interval for $\hat{\alpha}$. If your computer is old or your optimization algorithm is slow, you will not be able to take large M.

The data set for this problem, consisting of 20 x values, is posted on the course web site.