Christiano
FINC 520, Spring 2008
Final Exam.

This is a closed book exam. Points associated with each question are provided in parentheses. Good luck!

1. (20) Consider a stochastic process, a_t, which is the sum of two stochastic processes, a_t^T, and a_t^P, with

$$\begin{aligned} a_t &= a_t^1 + a_t^2, \\ a_t^1 &= \lambda a_{t-1}^1 + \varepsilon_t^1 \\ a_t^2 &= \rho_1 a_{t-1}^2 + \rho_2 a_{t-2}^2 + \varepsilon_t^2, \end{aligned}$$

   where ε_t^1 and ε_t^2 are mean-zero processes that are iid across time, independent of each other, with variances σ_1^2 and σ_2^2, respectively. The variable, a_t, is observed, as is a signal about ε_t^2 :

$$S_t = \varepsilon_t^2 + \omega_t.$$

   Here, ω_t is a mean-zero measurement error, iid across time and independent of ε_t^1 and ε_t^2 at all leads and lags. Its variance is σ_ω^2. Set this process up in state-space, observer form. That is, specify ξ_t, F, Q, H, R and x_t

$$\begin{aligned} \xi_t &= F\xi_{t-1} + \varepsilon_t, \; E\varepsilon_t\varepsilon_t' = Q, \\ x_t &= H'\xi_t + \bar{\omega}_t, \; E\bar{\omega}_t\bar{\omega}_t' = R, \end{aligned}$$

   where x_t is the list of observed variables.

2. (20) Let $f(\theta)$ denote the prior distribution of the vector of parameters; let $f(y|\theta)$ denote the likelihood of the data, y, conditional on the values of the parameters, θ; let $f(y, \theta)$ denote the joint distribution of θ and y. Let θ_i denote the i^{th} element of θ, where $\theta = \begin{bmatrix} \theta_1 & \cdots & \theta_i & \cdots & \theta_n \end{bmatrix}$, and n is the number of elements in θ.

   - Derive Bayes' rule, a formula that relates $f(\theta|y)$, the posterior distribution of θ after observing y, to $f(y|\theta)$, $f(\theta)$ and the marginal distribution of y.

- Bayesians report what they learn from observing y by comparing $f(\theta|y)$ with $f(\theta)$. When θ has more than one element, this comparison turns out in practice to be a computationaly intensive exercise.

   (a) Consider first a possibility that is *not* usually considered in practice. In this approach one compares the graphs of $f(\theta|y)$ and $f(\theta)$, allowing θ_i to vary and holding $\theta_{j\neq i}$ fixed (say, at the mode of $f(\theta|y)$). Explain the relatively minor computational complication that is involved in computing $f(\theta|y)$ for various values of θ_i, holding θ_j for $j \neq i$ fixed. Why is it that in practice, Bayesians do not display graphs like this?

   (b) Provide a formal definition of the posterior marginal distribution of θ_i. Why is it that Bayesians prefer to compare the marginal posterior distribution of θ_i with its marginal prior, over the object considered in (a)?

   (c) In practice, prior distributions over the different elements of θ are assumed to be independent. Provide an example to illustrate why this assumption might be questionable in some cases. Show why the computation of the marginal prior distribution of θ_i is trivial under the independence assumption.

3. (30) Consider a filter, $h(L)$, and let y_t be

$$y_t = h(L)\, x_t,$$

where x_t is a covariance stationary process with spectral density, $S_x(e^{-i\omega})$. Suppose that $h(L)$ is the 'band pass filter':

$$h\left(e^{-i\omega}\right) = \begin{cases} 1 & -b < \omega < -a,\ a < \omega < b \\ 0 & \text{otherwise.} \end{cases}.$$

It is said that "y_t is the output of a filter, $h(L)$, which passes through components of x_t with frequencies ω such that $h(e^{-i\omega}) = 1$ and shuts out all other frequencies".

   (a) Establish the result in the quote rigorously in the case of the following stochastic process for x_t:

$$x_t = \frac{\alpha}{2}\left[e^{-i\omega_1 t} + e^{i\omega_1 t}\right] + \frac{\beta}{2}\left[e^{-i\omega_2 t} + e^{i\omega_2 t}\right]$$

$$= \quad \alpha \cos(\omega_1 t) + \beta \cos(\omega_2 t),$$

where $h(e^{-i\omega_1}) = 1$ and $h(e^{-i\omega_2}) = 0$. In addition, α and β are two independent random variables with $E\alpha = E\beta = 0$, $E\alpha^2 = \sigma_\alpha^2$, $E\beta^2 = \sigma_\beta^2$.

(b) Does the x_t process in (a) satisfy covariance stationarity? Explain.

(c) Derive a formula for h_j, $j = 0, \pm 1, \pm 2, ...$ where

$$h(L) = \sum_{j=-\infty}^{\infty} h_j L^j.$$

4. (30) Shorter questions:

(a) prove that $E[x] = E\{E[x|y]\}$.

(b) provide the intuition underlying the Kalman 'updating formula' that pertains to the state-space/observer system in question (1):

$$\xi_{t|t} = \xi_{t|t-1} + P\left[\xi_t - \xi_{t|t-1} | x_t - H'\xi_{t|t-1}\right].$$

Here, $P[x|\Omega]$ denotes the projection of x onto the space composed of linear combinations of the elements in Ω. Also, $\xi_{t|t-j} \equiv P[\xi_t | x_{t-j}, x_{t-j-1}, ..., x_1]$, for $j \geq 0$.

(c) Derive an expression for the Kalman gain matrix, K_t

$$P\left[\xi_t - \xi_{t|t-1} | x_t - H'\xi_{t|t-1}\right] = K_t\left[x_t - H'\xi_{t|t-1}\right].$$

You may use a result concerning a necessary and sufficient condition for projections. But, if you use it, you must state it precisely.

(d) Prove that $\xi_{t|t-1} = F\xi_{t-1|t-1}$.

(e) Use the preceding results to derive an explicit expression for the projection,

$$P[\xi_t | x_t, x_{t-1}, ...].$$

(f) Provide a counterexample to the proposition that convergence in probability implies convergence in mean square.

(g) Consider the following ARCH process:

$$y_t = \varepsilon_t \sqrt{\alpha_0 + \alpha_1 y_{t-1}^2},$$

where $\alpha_i > 0$, $i = 1, 2$ and ε_t is iid over time and independent of y_{t-j}, $j > 0$. Define a martingale difference sequence (m.d.s.) and show that y_t is a m.d.s.