Problem Set VI (due Wed 5/20/09)

1. Let \( \{y_t\} \) be a stationary AR(2) process generated by the difference equation
   \[
   (1 - \alpha B)(1 - \beta B)y_t = \varepsilon_t,
   \]
   where the \( \{\varepsilon_t\} \) are white noise. Assume \( \alpha \neq \beta \) and both parameters are less than one in absolute value. Find the coefficients \( c_j \) such that \( y_t \) has the representation
   \[
   y_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}.
   \]

2. (a) Given two observations \( y_1 \) and \( y_2 \) from the causal AR(1) process
   \[
   Y_t - \alpha Y_{t-1} = U_t, \quad \{U_t\} \sim iid N(0, \sigma^2),
   \]
   such that \(|y_1| \neq |y_2|\), find the maximum likelihood estimates of \( \alpha \) and \( \sigma^2 \).

   (b) Let \( \{X_t\} \) be a stationary process with mean \( \mu \). Show that the optimal linear forecast of \( X_{n+h} \) based on \( 1, X_1, \ldots, X_n \) equals \( \mu \) plus the optimal linear forecast of \( Y_{n+h} \) based on \( Y_1, \ldots, Y_n \), where the zero mean process \( \{Y_t\} \) is defined by \( Y_t := X_t - \mu \) and where \( h \) is a fixed integer \( \geq 1 \).

3. Assume the model is given by
   \[
   \begin{align*}
   Y_t &= \mu + \alpha Y_{t-1} + U_t \text{ for } t = 1, \ldots, T, \\
   U_t &\equiv iid N(0, \sigma^2) \text{ for } \sigma^2 > 0, \\
   Y_0 &\sim N(\mu, \sigma^2/(1 - \alpha^2)).
   \end{align*}
   \]
   For \( T = 100, \mu = .5, \alpha = 0, .3, .6, .9, .99 \) and \( \sigma^2 = 1 \) generate \( R = 2000 \) time series according to this model. Each time, calculate the OLS estimator \( \hat{\alpha} \) of \( \alpha \) (from a regression of \( Y_t \) on a constant and \( Y_{t-1} \)). For the five simulation designs, report mean and median bias, standard deviation and RMSE (root mean squared error) of \( \hat{\alpha} \). Discuss your findings on the finite-sample behavior of the OLS estimator. Calculate nominal 95% confidence intervals (CI) for \( \alpha \) obtained from inverting a t-test. Report the empirical coverage probability of the CI. Discuss your findings.

4. Suppose that an MA(2) model is estimated when the second moving average parameter, \( \theta_2 \), is actually equal to zero. Find an expression for the relative efficiency of the resulting estimator of \( \theta_1 \) as compared with the estimator obtained from an MA(1) model.

Hints:
Equation (5.5.4) in Hamilton for the conditional likelihood of an MA(q) process (conditioned on \( \varepsilon_0 = 0, \ldots, \varepsilon_{-q+1} = 0 \), assuming \( \mu = 0 \) for simplicity) reads

\[
\mathcal{L} = \text{const} - \frac{T}{2} \log \sigma^2 - \frac{1}{2\sigma^2} S(\theta), \quad \text{for}
\]

\[
S(\theta) : = \sum_{t=1}^{T} \varepsilon(t)^2,
\]

\[
\varepsilon(t) : = y_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q},
\]

\[
\theta : = (\theta_1, \ldots, \theta_q).
\]

Leaving out the argument \( \theta \) in \( \varepsilon(t) \) (to simplify notation) we have

\[
\frac{\partial \mathcal{L}}{\partial \theta} = -\frac{1}{\sigma^2} \sum_{t=1}^{T} (\partial \varepsilon_t / \partial \theta) \varepsilon_t = \frac{1}{\sigma^2} \sum_{t=1}^{T} x_t \varepsilon_t \in R^{1\times q}, \quad \text{where}
\]

\[
x_t : = -(\partial \varepsilon_t / \partial \theta).
\]

From ML theory we know how to express the asymptotic variance of the ML estimators in terms of the information matrix

\[
E(\frac{\partial \mathcal{L}}{\partial \theta})'(\frac{\partial \mathcal{L}}{\partial \theta}) = E\left(\frac{1}{\sigma^2} \sum_{t=1}^{T} x_t \varepsilon_t\right)' \frac{1}{\sigma^2} \sum_{t=1}^{T} x_t \varepsilon_t
\]

\[
= \frac{1}{\sigma^4} \sum_{t=1}^{T} E x'_t x_t \varepsilon_t^2
\]

\[
= E \frac{1}{\sigma^2} \sum_{t=1}^{T} x'_t x_t.
\]

In fact, the asymptotic variance of the MLE estimator is the limit of \( T \)-times the inverse of this matrix, \( \lim_{T \to \infty} T (E \frac{1}{\sigma^2} \sum_{t=1}^{T} x'_t x_t)^{-1} \) or the inverse of the probability limit of \( \frac{1}{\sigma^2} \frac{1}{T} \sum_{t=1}^{T} x'_t x_t \).

**The case of MA(1):**

By (1), we have

\[
(\partial \varepsilon_t / \partial \theta) = -\theta_1 (\partial \varepsilon_{t-1} / \partial \theta) - \varepsilon_{t-1}
\]

and thus it follows that \( x_t \) follows an AR(1):

\[
x_t = -\theta_1 x_{t-1} + \varepsilon_{t-1}.
\]

Therefore

\[
p \lim T^{-1} \sum_{t=1}^{T} x_t x'_t = var x_t = \sigma^2/(1 - \theta_1^2),
\]

the formula for the variance of an AR(1). Therefore, using the above information matrix formula, it follows that the asymptotic variance of the conditional MLE of \( \theta_1 \) is given by

\[
(1 - \theta_1^2).
\]

Generalize this approach to MA(2) and you are done.