Econometrics, second half, second problem set. Turn this in at the final exam.

1. Estimate a logit model by GMM using the 10 independent data points

y	0	0	0	1	1	1	1	1	1	1
x	-1	-1	1	0	-1	-1	1	1	2	2

   For the logit model, the probability $P(y_t = 1|x_t) = (1 + \exp(-\theta_1 - \theta_2 x_t))^{-1}$, and the probability that $y_t = 0$ is the complement.

   (a) create a data file that contains these observations

   (b) find the conditional mean $E(y|x)$ and the conditional variance $V(y|x)$

   (c) propose at least 2 moment conditions, using the mean and the variance you found in (b)

   (d) write a Matlab function that computes the GMM estimator using your two moment conditions

   (e) compute the two step efficient GMM estimator

   (f) comment on the results

2. Prove that the GMM estimator based upon the g moment conditions $m_n(\theta) = \left[\begin{array}{cc} p_n'(\theta) & q_n'(\theta) \end{array}\right]'$ and the corresponding true optimal weight matrix is asymptotically more efficient than the GMM estimator based upon the $h < g$ moment conditions $p_n(\theta)$ and the corresponding true optimal weight matrix.

   (a) Interpret the result

   (b) Discuss the importance of the result from an empirical point of view. Are there any cautions one should observe when doing applied GMM work? Describe any problems you can imagine.

3. Suppose we have two equations

$$\begin{aligned} y_{t1} &= \alpha_1 + \alpha_2 y_{t2} + \epsilon_{t1} \\ y_{t2} &= \beta_1 + \beta_2 x_t + \epsilon_{t2} \end{aligned}$$

   where $V(\epsilon_{t1}) = \sigma_1^2 > 0$, $V(\epsilon_{t2}) = \sigma_2^2 > 0$, $E(\epsilon_{t1}\epsilon_{t2}) = \sigma_{12} \neq 0$. The observations are independent over time. The variable x_t is strictly exogenous: it is uncorrelated with the two epsilons at all time periods.

   (a) Is the OLS estimator of the parameters of the first equation consistent or not? Explain.

   (b) Is the OLS estimator of the parameters of the second equation consistent or not? Explain.

   (c) If the OLS estimator of the parameters of the first equation is not consistent, propose a consistent estimator of the parameters of the first equation and explain why the proposed estimator is consistent.

(d) If the OLS estimator of the parameters of the second equation is not consistent, propose a consistent estimator of the parameters of the second equation and explain why the proposed estimator is consistent.

4. Given the 10 independent data points

y	0	0	0	1	1	1	2	2	2	3
x	-1	-1	1	0	-1	-1	1	1	2	2

For the Poisson model, the density $f_Y(y|x) = \frac{\exp(-\lambda)\lambda^y}{y!}$, $y = 0, 1, 2, ...$ To make the model depend on conditioning variables, use the parameterization $\lambda(x) = \exp(\theta_1 + \theta_2 x)$.

   (a) the mean of a Poisson distribution with parameter λ is equal to λ, and so is the variance. Propose moment conditions to an overidentified $(g > k)$ GMM estimator of θ_1 and θ_2.

   (b) Estimate the parameters using two-step efficient GMM, using the moment conditions you have proposed.

   (c) discuss the results, and compare them to your ML estimates from the first problem set.

5. Write a Matlab script that generates two independent random walks, $x_t = x_{t-1} + u_t$ and $y_t = y_{t-1} + u_t$, where the initial conditions are $x_0 = 0$ and $y_0 = 0$, and the two errors are both iid N(0,1). Use a sample size of 1000: $t = 1, 2, ..., 1000$.

   (a) regress y upon x and a constant.

   (b) discuss your findings, especially the slope coefficient, the t statistic of the slope, and R^2. Are the finding sensible, given that we know that x has nothing to do with y?

   (c) compute the variance of y_t and x_t conditional on the initial conditions $y_0 = 0$ and $x_0 = 0$. Does the variance depend on t?

   (d) which of the assumptions of the classical linear regression model are not satisfied by this data generating process?