# Lecture 3: Generalized Method of Moments

Chris Conlon

February 7, 2023

NYU Stern

# Testing

# Testing with MLE

Before we discuss testing under GMM, let's look at testing under MLE.

- Helpful to define the likelihood ratio

$$LR \equiv -2 \cdot \ln \left[ \frac{\mathcal{L}(\theta_1|x)}{\mathcal{L}(\theta_2|x)} \right] = -2 \cdot [\ell(\theta_1|x) - \ell(\theta_2|x)]$$

- Consider $\dim(\theta_1) = q_1$ and $\dim(\theta_2) = q_2$ number of parameters
- Often we let $\theta_2$ be the unrestricted and $\theta_1$ be the restricted model.
- Define the degrees of degrees of freedom $N - \dim(\theta)$.
- The $LR$ statistic is distributed:

$$\Lambda \sim \chi^2_{q_1 - q_2}$$

- If we know $\theta_1$ and $\theta_2$ and we fix significance level $\alpha = 0.05$ then Neyman-Pearson Lemma says this is uniformly post powerful test

# Testing with MLE

We can consider the more advanced possiblity:

$$LR = -2 \ln \left[ \frac{\sup_{\theta \in \Theta_1} \mathcal{L}(\theta)}{\sup_{\theta \in \Theta} \mathcal{L}(\theta)} \right]$$

- $\Theta_1$ is a restricted version of the larger set $\Theta$.
- We can also consider non nested tests by looking at differences in degrees of freedom
  - This is mostly beyond what we will do in this course.
  - But we could ask: is $x_i$ distributed normally? or log-normally?

# GMM: J-test

The equivalent test in GMM is the J-test

$$Q_N(\theta) = g_N(\theta)' W_N g_N(\theta)$$
$$N \cdot Q_N(\theta) \longrightarrow^D \chi^2_{n-k}$$

This is an $LR$-type test statistic.

## Inverting LR tests

A useful technique is that we can always invert a test statistic in order to construct confidence intervals.

- Form an unrestricted estimate $\widehat{\theta}_{MLE}$ or $\widehat{\theta}_{GMM}$
- Compute $\ell(\widehat{\theta})$.
- Find all of the $\theta$ such that $CI = \{\theta : \ell(\widehat{\theta}) - \ell(\theta) < c\}$.
- If we do *GMM* we can use the *J*-stat instead.

How to choose $c$ the critical value.

- Compute the number of degrees of freedom / additional restrictions
- Choose a significance level $\alpha$ (ie: $\alpha = 0.05$).

# Confidence Intervals and Wald Tests

The multivariate Wald Test is:

$$H_0 : R\theta = r \quad H_1 : R\theta \neq r$$

$$\left(R\hat{\theta}_n - r\right)' \left[R\left(\hat{V}_n/n\right)R'\right]^{-1}\left(R\hat{\theta}_n - r\right) \quad \rightarrow \quad \chi^2_q$$

- $R$ is a matrix of $q$ linear restrictions on $k$ parameters.
- $\hat{V}_n$ is the covariance matrix for $\widehat{\theta}$.

You've been constructing CI's this way already

$$\widehat{\beta} \pm 1.96 SE(\widehat{\beta})$$

# LM or Score Test

There is a third test known as the Score Test or LM Test

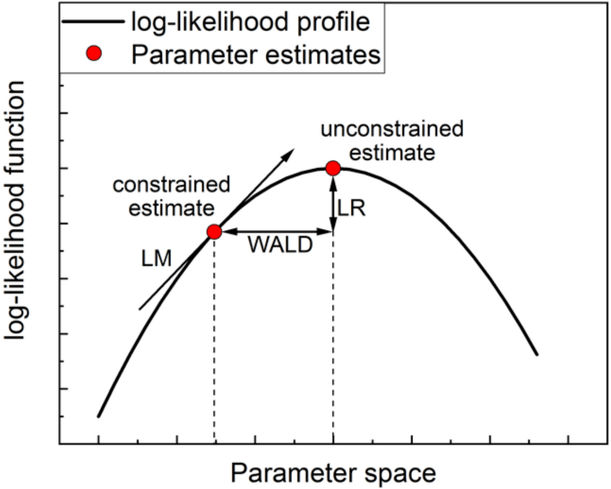
$$S(\theta) = \frac{\partial \ell(\theta|x)}{\partial \theta}$$
$$I(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \ell(X;\theta)|\theta\right]$$

- Compute the score of log likelihood
- Compute the Fisher Information.
- The test statistic

$$S^T\left(\hat{\theta}_0\right) I^{-1}\left(\hat{\theta}_0\right) S\left(\hat{\theta}_0\right) \sim \chi^2_q$$

Where $q$ is number of restrictions and $\theta_0$ is the true value.

# The Trinity of Testing



# What to do in practice?

- By reporting asymptotic standard errors you are implicitly using Wald type statisitics.
- If you are comparing models, you should probably try an LR type statistic if you can.
  - It used to be people didn't do this because *LR* required maximizing the objective function more than once.
  - But computers today are pretty good...
- For most extremum estimators (MLE, GMM, GEL, etc.) there are all three kinds of test-statistics
  - ... and around the true $\theta_0$ as $N \longrightarrow \infty$ they should coincide.
  - but in finite sample... anything can happen!

# Thanks!