OSM Boot Camp **Econ Notes**

Harrison Beard

Summer 2018

- ⓘ **Note.** Recall

$$\hat{\theta}_{\text{MLE}} = \theta : \max_{\theta} \log \mathcal{L}\left(\mathbf{x} \mid \theta\right)$$

$$\begin{aligned}\hat{\theta}_{\text{GMM}} &= \theta : \min_{\theta} \|m(\mathbf{x} \mid \theta) - m(\mathbf{x})\| \\ &= \arg\min_{\boldsymbol{\theta}} e(\mathbf{x} \mid \theta)^{\mathsf{T}} \mathbf{W} e(\mathbf{x} \mid \theta).\end{aligned}$$

where we can set

$$m_1(\mathbf{x} \mid \mu, \sigma) = \begin{pmatrix} \mathrm{E}\left[\mathbf{x} \mid \mu, \sigma\right] \\ \mathrm{Var}\left[\mathbf{x} \mid \mu, \sigma\right] \end{pmatrix}$$

and

$$m_2\left(\mathbf{x} \mid \mu, \sigma\right)$$

- **Topic**. OLS.
  - Consider

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i,$$

where

$$\mathrm{E}\left[\varepsilon_i\right] = 0$$

and

$$\mathrm{E}\left[x_{ji}\varepsilon_i\right] = 0.$$

  - Let

$$\hat{\theta}_{\text{GMM}} = \theta : \min_{\theta} \varepsilon^{\mathsf{T}} \varepsilon.$$

This is just OLS regression, in the form of GMM.

  - Now consider

$$\mathrm{E}\left[x_{1i}\varepsilon_i\right] = 0$$

and

$$\mathrm{E}\left[x_{2i}\varepsilon_i\right] = 0.$$

  - Computing a moment:

$$\varepsilon_i = y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}.$$

Each error is a function of the data.

  - Let

$$\begin{aligned}m_1\left(\mathbf{x} \mid \beta_0, \beta_1, \beta_2\right) &= \frac{1}{N}\sum_{i=1}^{N} \\ &= \frac{1}{N}\sum_{i}\left(y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}\right) \\ &= 0.\end{aligned}$$

MLE says that choose a set of params so that that sum adds up to zero.

$$\begin{aligned} m_2(\mathbf{x} \mid \beta_0, \beta_1, \beta_2) &= \frac{1}{N}\sum_{i=1}^{N} x_{1i}(y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}) \\ &= 0. \end{aligned}$$

This is another moment condition.

$$m_3(\cdot) = \cdots x_{2i} \cdots$$

and so on.

- **Topic**. Brock-Mirman model.

$$1 = \frac{\beta \mathrm{E}[r_{t+1} u'(c_{t+1})]}{u'(c_t)}$$

$$\implies \beta \mathrm{E}\left[\frac{r_{t+1} c_t}{c_{t+1}}\right] - 1 = 0.$$

Middle term is

$$\beta \mathrm{E}\left[\frac{\alpha \mathrm{e}^{z_{t+1}} k_{t+1}^{\alpha-1} c_t}{c_{t+1}}\right] - 1 = 0.$$

This is equation (9) in the notes.

- **Topic**.
  - Consider the difference between

$$\hat{\boldsymbol{\theta}}_{\mathrm{MLE}} = \arg\max \boldsymbol{\theta} \log \mathcal{L}(\mathbf{x} \mid \boldsymbol{\theta})$$

and

$$\hat{\boldsymbol{\theta}}_{\mathrm{GMM}} = \arg\min_{\boldsymbol{\theta}} \left\| \underbrace{m(\mathbf{x} \mid \boldsymbol{\theta})}_{\text{model}} - \underbrace{m(\mathbf{x})}_{\text{data}} \right\|$$

- **Topic**.
  - Let

$$S \coloneqq \# \text{ of sims } (s)$$

$$\tilde{\mathbf{x}} \coloneqq \{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_S\} \implies \tilde{\mathbf{x}}_s = \begin{pmatrix} y_{1s} x_{11s} x_{21s} \\ \vdots \\ y_{is} x_{1is} x_{2is} \\ \vdots \\ y_{NS} x_{1NS} x_{2NS} \end{pmatrix}.$$

The model moments are

$$m(\tilde{\mathbf{x}} \mid \boldsymbol{\theta}) = \frac{1}{S}\sum_{s=1}^{S} m(\tilde{\mathbf{x}}_s \mid \boldsymbol{\theta}).$$

Note that S is usually a large number, like $10,000$.

- **Topic**.
  - For SMM, we have

$$\hat{\boldsymbol{\theta}}_{\mathrm{SMM}} = \arg\min_{\boldsymbol{\theta}} \| m(\tilde{\mathbf{x}} \mid \boldsymbol{\theta}) - m(\mathbf{x}) \|.$$

In the L^2 norm way,

$$\arg\min_{\boldsymbol{\theta}} e(\tilde{\mathbf{x}} \mid \boldsymbol{\theta})^{\mathsf{T}} \mathbf{W} e(\tilde{\mathbf{x}} \mid \boldsymbol{\theta}),$$

where

$$e(\tilde{\mathbf{x}} \mid \boldsymbol{\theta}) \coloneqq m(\tilde{\mathbf{x}} \mid \boldsymbol{\theta}) - m(\mathbf{x}).$$

- **Topic**.

- Taking draws from the truncated normal distribution.

  Let the PDF be $\phi(\mathbf{x} \mid \boldsymbol{\theta})$ and the CDF be $\Phi(\mathbf{x} \mid \boldsymbol{\theta})$. To simulate a general distribution, here are the steps:

  1. Draw N values $u_i \sim \mathrm{Unif}(0, 1)$.
  2. Use $\Phi(\mathbf{x} \mid \boldsymbol{\theta})$ to convert u_i to x_i (the implied values from this PDF) $\implies x_i \sim \phi(\mathbf{x} \mid \boldsymbol{\theta})$.

  ! **Note.** The SMM problem will be a bonus problem.

Mon, 23 Jul. 2018

- **Topic**. **Lucas Tree Model**.
- ⓘ **Note.** Review of probability.
  - Start with probability space $Z \coloneqq \{z_1, z_2, \ldots, z_n\}$. Assume Z stays constant over time. Each z_i is mutually exclusive and exactly one must occur.
  - Take the infinite cartesian product of this set, Z^∞. We are interested in an infinite horizon. Call this set Ω.
  - For $\omega \in \Omega$, we have $\omega = \left(z^1, z^2, z^3, \ldots, z^t, \ldots\right)$. Call this a *path*.
  - Random variable $X(\omega) : \Omega \to \mathbb{R}$.
- ◊ **Example.** Simple random variables.
  - Suppose $z^1 = z_3$. Then

$$\begin{aligned} X(\omega) &= \left\{a \middle| z^1 = a\right\} \\ &= z_3. \end{aligned}$$

    Alternatively, let

$$X'(\omega) = \left\{a \middle| z^2 = a\right\}.$$

    Consider two different paths:

$$(z_1, z_2 z_7, z_{14}, z_2, z_1, z_{25}, \ldots)$$

    and

$$(z_1, z_7, z_{14}, z_2, z_{27}, \ldots).$$

    These paths are isomorphic to decision/probability tree paths.
- △ **Def.** A **filtration** is a sequence of σ-algebras or partitions.
- ◊ **Example.** Consider $Z = \{1, 2\}$. Two possible paths are

$$(2, 1, 2, 2, 1, 1, 1, 2, 1, \ldots)$$

  and

$$(1, 2, 2, 2, 1, 1, 2, 1, \ldots).$$

  - We can partition this set into two possible sets:

$$\{\omega | z_1 = 1\}$$

    and

$$\{\omega | z_1 = 2\}.$$

    If we combined these two sets with $\emptyset$ and Ω, we get a σ-algebra! Call this **S.A. 1.**
  - Now consider **S.A. 2.** What could it be? We have

$$\{\omega | z_1 = 1, z_2 = 1\}, \cdots, \{\omega | z_1 = 2, z_2 = 2\};$$

    in total we have four sets. Now include $\emptyset$ and Ω. If we include all the unions and cross-unions, we would also get a σ-algebra. In other words, this would be the σ-algebra *generated* by these sets. Note that this resultant σ-algebra is **finer** than S.A. 1. In other words, S.A. 1 is **coarser** than S.A. 2.
  - We make a sequence of successively finer σ-algebras, as we learn more information. This would be called a **filtration**.
  - Let $\mathrm{P}\{1\} =: \pi(1) = \frac{1}{3}$ and $\pi(2) = \frac{2}{3}$, and the outcomes IID.
- **Topic**. Consider the investor's problem,

$$\max_{\mathbf{c}} \left\{ \mathrm{E}\left[\sum_{t=0}^{\infty} \beta^t u(c_t)\right]\right\}.$$

Let's say we have

Asset	Price	Dividend
Asset 1	p_t	$d_t(\omega)$,

where the price and the dividend are both random variables. The only things that we can consume are with the dividends that are paid out by these assets.

- Define $\boldsymbol{\theta}_t$ as the household's portfolio at time t. This might be

$$\boldsymbol{\theta}_t = (\theta_{1t}, \theta_{2t}, \ldots, \theta_{nt}),$$

  i.e., how many shares of each stock you have in your portfolio.
- Households take their wealth at time t and choose to eat some of it and invest the rest of it. At time t,

$$c_t + p_{1t}\theta_{1t} + p_{2t}\theta_{2t} + \cdots + p_{nt}\theta_{nt}.$$

  This represents how much the agent consumes and distributes investment across different assets at time t. Note that

$$c_t + p_{1t}\theta_{1t} + p_{2t}\theta_{2t} + \cdots + p_{nt}\theta_{nt} \le (p_{1t} + d_{1t})\,\theta_{1,t-1} + \cdots + (p_{nt} + d_{nt})\,\theta_{n,t-1}.$$

  In vector notation,

$$\underbrace{c_t}_{\text{not a vector}} + \mathbf{p}_t'\boldsymbol{\theta}_t \le (\mathbf{p}_t' + \mathbf{d}_t')\,\boldsymbol{\theta}_{t-1}.$$

  Rewriting, we have

$$\begin{aligned} c_t &= \mathbf{p}_t'\boldsymbol{\theta}_{t-1} + \mathbf{d}_t'\boldsymbol{\theta}_{t-1} + \mathbf{p}_t'\boldsymbol{\theta}_t \\ &= \mathbf{p}_t'(\boldsymbol{\theta}_{t-1} - \boldsymbol{\theta}_t) + \mathbf{d}_t'\boldsymbol{\theta}_{t-1}. \end{aligned}$$

  So our problem is

$$\max_{\boldsymbol{\theta}} \left\{ \mathrm{E}\left[\sum_{t=0}^{\infty} \beta^t u\Big(\underbrace{\mathbf{p}_t'(\boldsymbol{\theta}_{t-1} - \boldsymbol{\theta}_t)}_{\substack{\uparrow \\ \textit{change in portfolio wealth}}} + \mathbf{d}_t'\boldsymbol{\theta}_{t-1} \Big) \right] \right\}.$$

- So what is this individual going to do? Take the FOCs of the above with respect to $\boldsymbol{\theta}_t$:

$$\mathrm{E}\left[\beta^t \Big(-u'(\underbrace{c_t}_{\substack{\uparrow \\ \mathbf{p}_t'\boldsymbol{\theta}_{t-1} + \mathbf{d}_t'\boldsymbol{\theta}_{t-1} + \mathbf{p}_t'\boldsymbol{\theta}_t}})\mathbf{p}_t + \beta u'(c_{t+1})(\mathbf{p}_{t+1} + \mathbf{d}_{t+1}) \Big) \right] = \mathbf{0}^{\mathsf{T}}.$$

  For row i:

$$\mathrm{E}\left[\beta^t \left(-u'(c_t)\,p_{it} + \beta u'(c_{t+1})\,(p_{i,t+1} + d_{i,t+1}) \right) \right] = 0.$$

  Divide the LHS and the RHS by p_{it}:

$$\mathrm{E}\left[\beta^t \left(-u'(c_t) + \beta u'(c_{t+1}) \underbrace{\frac{p_{i,t+1} + d_{i,t+1}}{p_{it}}}_{R_{it}} \right) \right] = 0.$$

Wed, 25 Jul. 18

Stochastic Discount Factor.

- $$\begin{aligned}
\mathrm{E}\left[-\beta^t u'(c_t)\,\mathbf{p}_t + \beta^{t+1} u'(c_{t+1})(\mathbf{p}_{t+1}+\mathbf{d}_{t+1})\right] &= 0 \\
\mathrm{E}\left[-u'(c_t)\,\mathbf{p}_t + \beta u'(c_{t+1})(\mathbf{p}_{t+1}+\mathbf{d}_{t+1})\right] &= 0 \\
\implies -u'(c_t) + \beta \mathrm{E}\left[u'(c_{t+1}) R_{it} \middle| \Omega_t\right] &= 0 \\
\implies u'(c_t) &= \beta \mathrm{E}\left[u'(c_{t+1}) R_{it} \middle| \Omega_t\right].
\end{aligned}$$

  $$1 = \mathrm{E}\left[\underbrace{\boxed{\beta \frac{u'(c_{t+1})}{u'(c_t)}}}_{\text{stochastic discount factor}} R_{it} \middle| \Omega_t\right].$$

  Note that for small agents, R_{it} is exogenous to choices. Also,

  $$\begin{aligned}
1 &= \mathrm{E}\left[m_t R_{it} \middle| \Omega_t\right] \\
1 &= \mathrm{Cov}\left[m_t, R_{it} \middle| \Omega_t\right] + \mathrm{E}\left[m_t \middle| \Omega_t\right] \mathrm{E}\left[R_{it} \middle| \Omega_t\right].
\end{aligned}$$

  For now, write $\mathrm{Cov}\left[m_t, R_{it} \middle| \Omega_t\right] =: \mathrm{Cov}\left[m_t, R_{it}\right]$ for brevity. We have

  $$\underbrace{\mathrm{Cov}\left[m_t, R_{it}\right] + \mathrm{E}\left[m_t\right]\mathrm{E}\left[R_{it}\right] = \mathrm{Cov}\left[m_t, R_{jt}\right] + \mathrm{E}\left[m_t\right]\mathrm{E}\left[R_{jt}\right]}_{\text{risk-return tradeoff}}$$

  (for $i \neq j$ in general).
- Recall

  $$m_t = \beta \frac{u'(c_{t+1})}{u'(c_t)}$$

  and

  $$R_{it} = \frac{\mathbf{p}_{i,t+1} + \mathbf{d}_{i,t+1}}{\mathbf{p}_{it}}.$$

ⓘ **Note.** Small investors do not have control over R; they take it as given, and $\mathrm{E}[R]$ is a belief based on information that the investor has. Portfolio choices $\boldsymbol{\theta}$ are encapsulated in m. As an investor, we are manipulation the $\mathrm{Cov}[m, R]$ term. In equilibrium, we all do that.

- consider the FOC

  $$-u'(c_t) + \beta \mathrm{E}\left[u'(c_{t+1})\left(\frac{1}{\mathbf{p}_{it}}\right)\right] = 0.$$

  Rearranging, we have

  $$\mathbf{p}_{it} = \beta \mathrm{E}\left[\frac{u'(c_{t+1})}{u'(c_t)}\right].$$

  This is the **expected discount factor**.

- **Gross return**:

$$\begin{aligned} R_{it} &= \underbrace{1 + r_{it}}_{(*)} \\ &= \frac{1}{\mathbf{p}_{it}} \\ &= \boxed{\frac{1}{\beta \mathrm{E}\left[\frac{u'(c_{t+1})}{u'(c_t)}\right]}}. \end{aligned}$$

  This is the implied rate of return by this model. This can be used to estimate the parameters of the SDF through GMM. This is a **moment condition**. Note that GMM was developed by Lars Hansen specifically to estimate these particular moment conditions. Also note that $(*)$ is something that we can easily observe. We can use all the expected marginal utilities information to estimate what people's β is.
  - We would assume that
  $$u(c_{t+1}) = \frac{c_{t+1}^{1-\gamma}}{1-\gamma}.$$
- If we take the FOC

$$\begin{aligned} g(\gamma, \beta, R) &= \beta \mathrm{E}\left[\frac{u'(c_{t+1})}{u'(c_t)} R_{it}\right] - 1 \\ &= 0 \end{aligned}$$

  as a *moment condition*. We will do this on the homework. Every period, calculate

$$\mathrm{E}\left[\frac{u'(c_{t+1})}{u'(c_t)} R_{it} - 1\right].$$

  Figure out what the average is, adjust the γ, and figure out what γ would make this condition equal to zero. (This is Questions 1 & 2.)

Kyle (1985) Model

- Agents:
  - **Market Makers**: The ones submitting limit orders. We want to understand this party. To these guys, V is a random variable, but they know the distribution of V. They don't know which traders are informed or uninformed. They *observe* $Y = X + U$ (the *sum* of informed and uninformed demand).
  - **Informed Traders**: The ones who actually *know* what V is exactly. They also know the distribution of V that market makers know. Their optimal demand is $X(V)$.
  - *Noise Traders* or **Uninformed/Liquidity Traders**: Individuals that are trading in ways that are uncorrelated with the future value of the asset. *There is no information in this trade*. No "market timing" effects involved. They know nothing. They could be bad traders, or they could just be laypeople selling stock in order to gain idiosyncratic personal financial liquidity. We assume $U \sim \mathcal{N}\left(0, \sigma_u^2\right)$. **Assume that U, V are uncorrelated.**
- Let future value of asset is V. By "future" for this model, we refer to a small time horizon, i.e., less than a week or so. Assume $V \sim \mathcal{N}(p_0, \Sigma_0)$. The market makers know this distribution.

Key Concept 1: Price Function.

△ **Def.** In **equilibrium**, a market maker sets a **price function** $P(Y)$, where $Y = X + U$. The informed traders also have a **demand** such that $\mathrm{E}\left[(V - P(Y))\underbrace{(-Y)}_{\text{position}} \middle| Y\right] = 0$ (competitive risk-neutral market makers), and where $X(U)$ maximizes $\mathrm{E}\left[(V - P(Y))X(U) \middle| V\right]$. *In general, this is a very hard system of equations to solve.*

- To solve for the equilibrium, we use *guess and check*. Assume that the equilibrium is $P(Y) = \mu + \lambda Y$. Assume linear. *Note that this is the place where you would expect to see linearity*—risk-neutral and linear-preference assumptions. Take this as given, and move directly to

$$\mathrm{E}\left[(V - P(Y))X(U) \middle| V\right].$$

They want to

$$\max_X \mathrm{E}\left[\left(V - \mu - \lambda\left(X + U\right)\right)X \middle| V\right].$$

Note that for informed traders, U is the only thing that is unknown. Now some simplification:

$$\begin{aligned} &= \mathrm{E}\left[VX - \mu X - \lambda X^2 - \lambda UX \middle| V\right] \\ &= VX - \mu X - \lambda X^2 - \lambda X \underbrace{\mathrm{E}\left[U \middle| V\right]}_{=0} \\ &= VX - \mu X - \lambda X^2. \end{aligned}$$

Calculating the FOC and solving for optimal X,

$$V - \mu - 2\lambda X = 0$$

so

$$\begin{aligned} X &= \frac{V}{2\lambda} - \frac{\mu}{2\lambda} \\ &= -\frac{\mu}{2\lambda} + \frac{1}{2\lambda}V. \end{aligned}$$

Note that the second derivative is negative if $\lambda > 0$. This is an important consideration.