Introduction to Macroeconomics

Rhys Bidder

Federal Reserve Bank of San Francisco

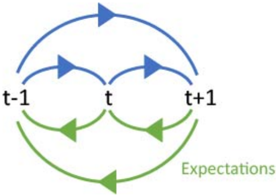
Michaelmas Term 2019

Disclaimer

The views expressed in this presentation, and all errors and omissions, should be regarded as those solely of the authors, and are not necessarily those of the Federal Reserve Bank of San Francisco, the Federal Reserve Board of Governors or the Federal Reserve System.

DSGE Models

Macroeconomics is Dynamic



Economic models are dynamic

- They connect 'yesterday' ($t-1$), 'today' (t) and 'tomorrow' ($t+1$)
- But so do other models (e.g. weather forecasting, bridge safety, ...)
- We share many tools with engineers, physicists...

Expectations are what make economics 'special'

- Our models (our world) are explicitly forward looking
- People make decisions with beliefs about the future in mind
- Aside: IoT, AI and machine learning?

Macroeconomics is Stochastic

Decision-makers must take randomness into account

- Economists use the word ‘stochastic’ to mean ‘randomness’

We often don’t know what’s going to happen

- We may not even know what has happened or is happening (filtering)

Risk: Even if we know how the world works there is still randomness

- Playing a *known* card game (don’t know what’s coming next but know the odds)

Uncertainty: There may also be ‘unknown unknowns’ - Rumsfeld (2002)

- Like playing an *unknown* card game (don’t know what’s coming next *and* don’t know the odds)

Macroeconomics is Stochastic

Impulses → Propagation → Fluctuations

Impulses: Unexpected 'structural shocks' to...

- Technology (ability to combine/reallocate resources effectively...)
- Monetary/fiscal policy (deviation from 'standard' procedure, wars,...)
- Oil price, Forex and terms of trade (esp. for 'small economies')

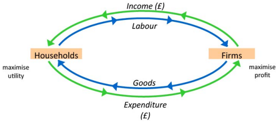
Propagation: How the shocks work their way through the system

- Decisions by households and firms
- Depends on preferences, technology, market structure, beliefs,...
- *Modern* macro: explicitly microfounded and tightly parameterized

Fluctuations: Due to shock propagation from current/previous periods

- Equilibrium of a model $\Rightarrow$ probability distribution over 'sequences'
- Shocks and their propagation together define this distribution

Macroeconomics is General Equilibrium



Markets are interconnected

- We allow spillovers and connections between markets
- Contrast with partial equilibrium (micro 101)
- Typically variables are 'simultaneously co-determined' in G.E.

Simultaneity $\Rightarrow$ difficult to talk about causal relationships

- Income depends on hours worked, depends on hiring, depends on scale of production, depends on demand, depends on income. . .

Interdependence of **individually optimal** decisions by multiple agents and **consistency** requirements in the aggregate

- $\Rightarrow$ Set of simultaneous equations

Macroeconomics is General Equilibrium

What does it mean to 'solve' an economic model?

- Models involve a lot of 'variables' (consumption, unemployment, output, wages,. . .)
- Accounting and technological constraints imply relationships among these variables
- The assumption that people and firms are optimizing also implies relationships among these variables
- There is a core set of variables that are needed to describe 'the current situation' (all the relevant info.)
- We call these variables **'the state'**
- **Solving a model $\Leftrightarrow$ finding functions that relate all the variables in the economy to the state**

Macroeconomics is General Equilibrium

Remember high school math, when you had to 'solve' systems of equations for x_1 and x_2

$$\alpha_{1,1}x_1 + \alpha_{2,1}x_2 = y_1 + z_1$$
$$\alpha_{2,1}x_1 + \alpha_{2,2}x_2 = y_2 + z_2$$

where the *solution* will be $x_1 = f_1(\alpha, y, z)$ and $x_2 = f_2(\alpha, y, z)$

DSGE models' optimality conditions (Factor price ratio=MRT for firms, say) and equilibrium requirements (resource constraints, market clearing, belief consistency) will also yield a system of equations

- More complicated (dynamics and expectations 'replace' y and z), but similar intuition
- Express endogenous variables as functions of the state
- The exact form of those functions will depend on the model's structure and parameters

Macroeconomics is General Equilibrium

In practice, questions in the press are frequently poorly posed

- **Q:** What happens to employment if inflation drops?
- **A:** It depends. Is inflation falling because of looser policy or because there's been an unexpected improvement in technology?

'Employment' and 'inflation' are **endogenous and co-determined**

- They may both be fluctuating because of other factors - rather than one causing the other
- Causality may work in both directions via response of different agents

In micro 101 (and in a lot of old skool macro) these issues are brushed aside

- Modern macro $\Rightarrow$ shocks are the exogenous 'causal' source of fluctuations
- Transmission mechanism (in equilibrium) then determines how variables co-move

D. . . S. . . GE

Multiple periods (D)

- Agents form dynamic plans taking future into account

Random shocks continually hitting the economy (S)

- From monetary policy and other sources (e.g. technology)

Optimization problems of individual agents (GE)

- Preferences/profit maximization with budget/technological constraints

Feasibility of optimal behavior in the aggregate (GE)

- Aggregation of individuals' decisions respects resource constraints

Consistency of beliefs with the induced path of the economy (GE)

- Typically impose Rational Expectations requirement

Stylized Facts

Macroeconomic fluctuations

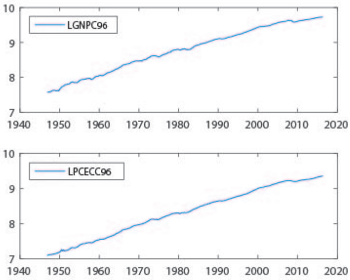
Decomposing fluctuations

- Underlying trend / low-frequency movements ('growth' literature)
- Business cycle
- Measurement error
- Seasonality (rarely discussed - let the statisticians handle it)
- 'Random' fluctuations ('unmodelable' - or stuff we've failed to model)

Growth and fluctuations can be connected as high frequency decisions can influence low frequency phenomena

- Consume/save/innovate $\rightarrow$ investment/technology $\rightarrow$ growth
- We will focus mainly on 'business cycle' fluctuations

Figure 1: Log GNP and consumption



- Note the fairly (until recently?) constant trend
- Long run co-movement between Y and C
- Fluctuations around trend with a few big hits (recessions)

Hodrick-Prescott Filter

Hodrick-Prescott filter decomposes y_t into trend, τ_t, and cycle, c_t

$$y_t = \tau_t + c_t$$

Trend should be 'smooth' but track data closely in longer run

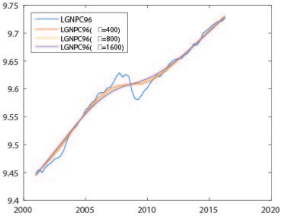
$$\min_{\{\tau_t\}_{t=1}^{T}} \left(\underbrace{\sum_{t=1}^{T} (y_t - \tau_t)^2}_{Tracking} + \lambda \underbrace{\sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2}_{Smoothness} \right)$$

λ is weighting parameter

- $\lambda = 0$ means $\tau_t = y_t$ and trend is data
- $\lambda \to \infty$ means $\Delta^2 \tau_t = 0$ and trend is linear

Hodrick-Prescott filter

Figure 2: Trend GNP using H-P filter



US GNP and H-P trends with different λ

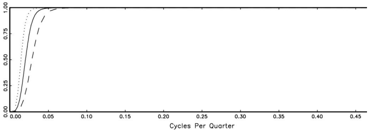
- See Harvey and Jaeger (1993) on (in)appropriate use of HP filter

There are other ways of extracting trends

- Linear (‘draw a line through it’)
- Quadratic or cubic (‘draw a bendy line through it’)
- Theoretical, rather than statistical (‘use a model’)

Hodrick-Prescott filter

Figure 3: PTF of H-P filter with $\lambda = 6400$ dots; 1600 solid; 400 dashes



The ‘power transfer function’ shows at what frequencies the filter allows variation (e.g. 0.05 cycles per quarter $\Rightarrow$ 5 years for a cycle)

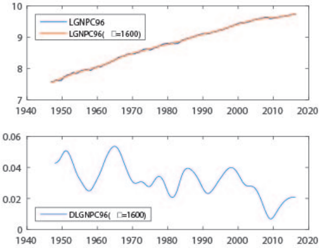
- Low frequencies: Few cycles per quarter (medium-long term trends)
- High frequencies: Many cycles per quarter (business cycles and noise)
- Variation in y_t is a mix of *all* frequencies - the filter lets through (‘passes’) only a subset

(Somewhat unconvincing) consensus that $\lambda = 1600$ distinguishes ‘business cycle’ and trend (quarterly data)

- See also Baxter and King (2006) ‘band-pass’ filter

Long run properties of macro data

Figure 4: GNP and Trend Growth Rate ($\lambda = 1600$)



Trend US growth 'traditionally' thought of as $\approx 2.5 - 3.0\%$

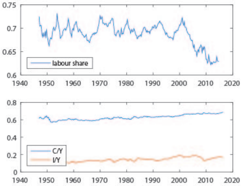
- GNP follows 'straight line' in logs

Note: Current debate about secular stagnation etc.

- In U.S. labor force growth slowing and effects of 'internet' subsiding
- See Fernald and Li (2019)

Long run properties of macro data

Figure 5: Labor share and 'Great Ratios'



Fraction of output that goes to labor approximately constant

- Or is it? Big debate currently about declining ratio.
- Change in bargaining power, skills biased technical change, declining competition, cheaper investment goods, off-shoring?

'Great Ratios' (consumption and investment shares) ≈ constant

- Maybe minor trend but these seem to be holding steady

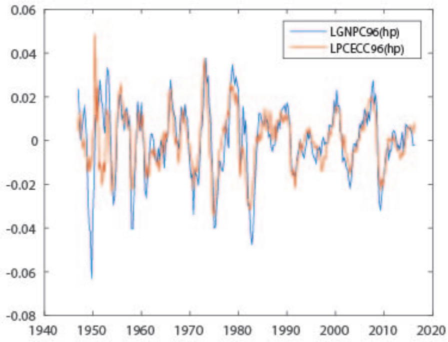
Kaldor (1957) facts

Kaldor's (1957) stylized facts have held up fairly well (subject to above caveats):

1. Output per worker grows at a roughly constant rate
2. Capital per worker grows over time
3. Capital/output ratio is roughly constant
4. Rate of return to capital is constant
5. Shares of capital and labor in net income are nearly constant
6. Real wage grows over time
7. Ratios of consumption and investment to GDP are constant

Short run properties of macroeconomic data

Figure 6: HP-Detrended GNP and Non-durable Consumption



- Both series are 'cyclical' components from HP filtering
- Strong positive correlation (consumption is 'pro-cyclical')
- Consumption smoother than output. Why?

Short run properties of macroeconomic data

Table 1: Cyclical behavior of US economy 1954 – 1991

Variable	Sd%	Cross-correlation of output with:						
		t-3	t-2	t-1	t	t+1	t+2	t+3
GNP	1.72	0.38	0.63	0.85	1.00	0.85	0.63	0.38
CND	0.86	0.55	0.68	0.78	0.77	0.64	0.47	0.27
CD	4.96	0.49	0.65	0.75	0.78	0.61	0.38	0.11
I	8.24	0.38	0.59	0.79	0.91	0.76	0.50	0.22
H	1.59	0.30	0.53	0.74	0.86	0.82	0.69	0.52
Ave H	0.63	0.34	0.48	0.63	0.62	0.52	0.37	0.23
L	1.14	0.23	0.46	0.69	0.85	0.86	0.76	0.59
GNP/H	0.90	0.20	0.30	0.33	0.41	0.19	0.00	-0.18
Ave W	0.55	0.21	0.14	0.09	0.03	-0.07	-0.09	-0.09

Stylized facts

- Non-durable consumption less volatile than output (consumption smoothing)
- Volatility of output and hours similar
- Employment more volatile than average hours (extensive margin, offset by intensive)
- Wages less volatile than productivity (smooth wages)
- Productivity slightly pro-cyclical
- Wage acyclical (despite employment volatility)

Vector Autoregressions

Vector autoregressions (VARs)

Example: p-order VAR in two variables (y_t and x_t)...

$$\begin{aligned} y_t &= a_{11}^{(1)} y_{t-1} + \ldots + a_{11}^{(p)} y_{t-p} + a_{12}^{(1)} x_{t-1} + \ldots + a_{12}^{(p)} x_{t-p} + u_{yt} \\ x_t &= a_{21}^{(1)} y_{t-1} + \ldots + a_{21}^{(p)} y_{t-p} + a_{22}^{(1)} x_{t-1} + \ldots + a_{22}^{(p)} x_{t-p} + u_{xt} \end{aligned}$$

where we imagine, say, y_t being a measure of real activity and x_t a measure of the stance of monetary policy (interest rate or money supply)

In matrix form...

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = A(L) \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_{yt} \\ u_{xt} \end{pmatrix}$$

where A(L) is a p-order matrix polynomial in the lag operator and u_{it} is an innovation ('forecast error') to variable i.

Vector Autoregressions

Consider the special case of a VAR(1) ($p = 1$)

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_{yt} \\ u_{xt} \end{pmatrix}$$

Thus we have

$$\begin{aligned} y_t &= a_{11} y_{t-1} + a_{12} x_{t-1} + u_{yt} \\ x_t &= a_{21} y_{t-1} + a_{22} x_{t-1} + u_{xt} \end{aligned}$$

Let us now consider the 'forecast errors', u_{yt} and u_{xt}...

Vector Autoregressions

x_t may not be what was expected at $t-1$ because

- the policymaker responded to surprises elsewhere in the economy in t
- the policymaker did something unexpected in a way unrelated to the broader economy (e.g. unexpected change in the voting patterns of a policy committee after new appointments)

We are interested in the effect of a *policy surprise* (i.e. originating with the policymaker) rather than a surprise to the policy instrument, *per se*

- **But** without further assumptions u_{xt} could be a combination of the desired 'policy shock' and an 'output shock' (such as a random loss in confidence by consumers, say)

Vector Autoregressions

x_t may not be what was expected at $t-1$ because

- the policymaker responded to surprises elsewhere in the economy in t
- the policymaker did something unexpected in a way unrelated to the broader economy (e.g. unexpected change in the voting patterns of a policy committee after new appointments)

We are interested in the effect of a *policy surprise* (i.e. originating with the policymaker) rather than a surprise to the policy instrument, *per se*

- **But** without further assumptions u_{xt} could be a combination of the desired 'policy shock' and an 'output shock' (such as a random loss in confidence by consumers, say)

We can think of the forecast errors as being a linear combination of 'structural' shocks, e_{yt} and e_{xt} where the latter is the 'policy shock'

$$\begin{pmatrix} u_{yt} \\ u_{xt} \end{pmatrix} = \begin{pmatrix} e_{yt} + \theta e_{xt} \\ \phi e_{yt} + e_{xt} \end{pmatrix} = \begin{pmatrix} 1 & \theta \\ \phi & 1 \end{pmatrix} \begin{pmatrix} e_{yt} \\ e_{xt} \end{pmatrix} \equiv B \begin{pmatrix} e_{yt} \\ e_{xt} \end{pmatrix}$$

Without assumptions on B (i.e. on θ and ϕ in this example) we cannot 'pull apart' the forecast errors (u_{it}) and separately **identify** the effect of the structural shocks (e_{it})

- Example 1: Assume $\phi = 0$
  - Policy variable does not respond contemporaneously to 'output shocks'
- Example 2: Assume $\theta = 0$
  - Policy only affects output with a lag
- Identified VAR is commonly referred to as a structural vector autoregression (SVAR)

Vector Autoregressions

Many identification schemes have been proposed (see Christiano *et al* (1999) and Ramey (2016))

- Two examples above are a case of using different 'Cholesky' orderings
- Ordering not unique and results may depend on ordering
- Should have a plausible story
- Logic extends to multiple shocks but more difficult to imagine a complete ordering

Some famous examples of identification based on ordering...

- Sims (1992)
- Christiano, Eichenbaum and Evans (2005) - only used one ordering assumption

Sims (1992)

Six variable VAR for UK 1965 – 1990 with causal ordering

1. Short interest rate (R)
2. Index of foreign exchange value of domestic currency
3. Commodity price index
4. Monetary aggregate
5. Consumer price index
6. Industrial production index

Innovations only affect variables (weakly) lower in causal ordering

- Interest rate first $\Rightarrow$ Central Bank not contemporaneously reacting to economic news

Sims (1992)

Figure 7: Sims (1992) VAR for U.K. Economy

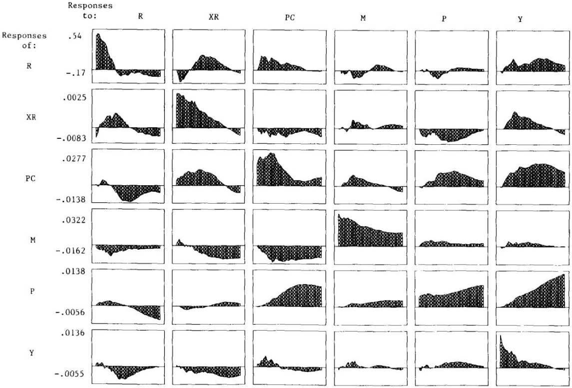


Fig. 4. United Kingdom, 1965:4–1990:12.

Christiano, Eichenbaum and Evans (2005)

Nine variable VAR for US 1965 – 1995 with causal ordering

1. Real GDP
2. Real consumption
3. GDP deflator
4. Real investment
5. Real wage
6. Labor productivity
7. Interest rate
8. Real profit
9. Growth rate of M2

Implications of position of R:

- R Shocks only affect real profit and M2
- R Affected by all shocks except real profit and M2 shocks

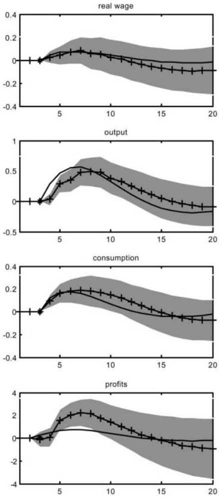
Christiano, Eichenbaum and Evans (2005)

Results suggest that after an expansionary monetary policy shock:

1. Output, consumption, and investment respond in a hump-shape, peaking after about one and a half years and returning to pre-shock levels after about three years
2. Inflation responds in a hump-shape, peaking after about two years
3. Interest rate falls for roughly one year
4. Real profits, real wages, and labor productivity rise
5. Growth rate of money rises immediately

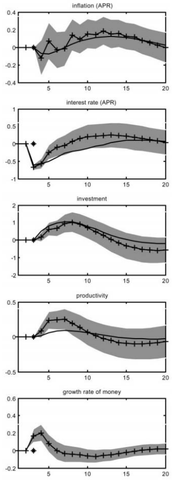
Christiano, Eichenbaum and Evans (2005)

Figure 8: CEE (2005) - Impulse responses (I)



Christiano, Eichenbaum and Evans (2005)

Figure 9: CEE (2005) - Impulse responses (II)



Problems with causal orderings

Sims (1992) and CEE (2005) find prices initially ↑ after unexpected R_t ↑

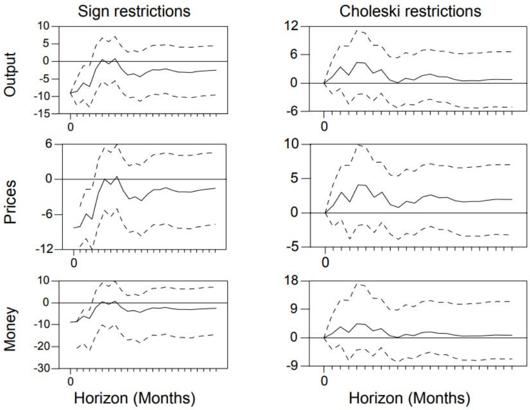
- Often referred to as the 'price puzzle'
- Could be a cost channel effect (firm borrowing costs rise so they increase prices) but more likely faulty identification
- Thoughts on other possible explanations?

Sign restriction VARs designed to rule out such anomalies

- Impose restrictions not by ordering but by restricting which direction variables should move in following a shock
- See Canova (2007), Canova and Paustian (2010), Uhlig (1998), Rubio-Ramirez (2018)
- Though if the problem is incorrect set of variables (information for forecast) then might not help?

Sign restrictions

Figure 10: Canova (2007) - Comparing responses to a monetary tightening based on Cholesky and Sign restrictions



Sign restrictions

Figure 11: Canova and Paustian (2010) - Sign restrictions

Table 2
Signs of the impact response intervals to shocks.

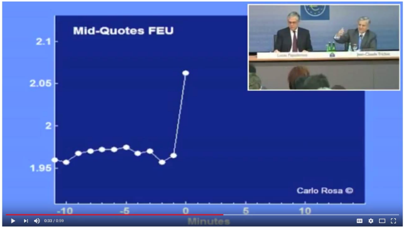
Variable	Markup shocks							Monetary shocks						
	M	M1	M2	M3	M4	N1	N2	M	**M1**	**M2**	M3	M4	N1	N2
R_t	+	+	+	+	+	+	+	+	+	+	+	+	+	+
w_t	–	–	–	–	–	–	–	?	+	–	?	?	?	?
π_t	+	+	+	+	+	+	+	–	–	–	–	–	–	–
y_t	–	–	–	–	–	–	–	–	–	–	–	–	–	–
n_t	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Taste shocks							Technology shocks						
	M	M1	M2	M3	M4	N1	N2	M	M1	M2	M3	M4	N1	N2
R_t	+	+	?	+	?	+	+	–	–	–	–	–	–	–
w_t	?	–	?	?	–	?	?	?	+	?	?	+	?	?
π_t	+	+	?	+	?	+	+	–	–	–	–	–	–	–
y_t	+	+	+	+	+	+	+	+	+	+	+	+	+	+
n_t	+	+	+	+	+	+	+	–	–	–	–	–	–	–

Identification using high frequency information

- ECB monthly press conference January 15, 2009
- Traders expect interest rate cut on February 5, 2009
- Trichet announces no policy change expected next meeting
- Traders revise up expectations

Rosa (2008)

Figure 12: Rosa (2008) - Mid-quote on 3-month Euribor future expiring in 03/09



Time zero when Trichet starts answering a journalist's question

- Presumption is that nothing else 'changes' during the short period from just before to just after $\Rightarrow$ Simply captures policy-surprise
- NOTE: Current debate about whether the CB is revealing info about the rest of the economy (as opposed to simply about monetary policy)

While researchers have disagreed on the best means of identifying policy shocks, there has been a surprising consensus on the general nature of the economic responses to monetary policy shocks. A variety of VARs estimated for a number of countries all indicate that, in response to a policy shocks, output follows a hump-shaped pattern in which the peak impact occurs several quarters after the initial shock.

- Walsh, 1998, p.31

Appendix: Some math(s)

Unofficial maths 'requirements'

Most of the maths we use will entail...

- Basic algebra
  - C_t will represent consumption in time t
- Basic probability
  - Mean/expectation and maybe standard deviation
- Collecting coefficients / factorization
  - $ax + bx = (a + b)x$
- Summations
  - $\sum_{j=0}^{J} f(x_j) \equiv f(x_0) + f(x_1) + \ldots + f(x_J)$
- Calculus
  - You will need to differentiate very simple functions
  - You will probably only need to understand what an integral ($\int$) *means*
  - You will need to be able to linearize and log-linearize

Solving an economic model

What does it mean to 'solve' an economic model?

- Models involve a lot of 'variables' (consumption, unemployment, output, wages,. . .)
- Accounting and technological constraints imply relationships among these variables
- The assumption that people and firms are optimizing also implies relationships among these variables
- There is a core set of variables that are needed to describe 'the current situation' (all the relevant info.)
- We call these variables 'the state'
- **Solving a model $\Leftrightarrow$ finding functions that relate all the variables in the economy to the state**

Taylor approximations

Consider consumption in time t, C_t

- A solved model implies

$$C_t = f(tax\ \ rate, income, assets, monetary\ \ policy, \ldots)$$

- Or let's just call the state, s_t

$$C_t = f(s_t)$$

- Sadly, it is rare that the function f can actually be calculated
- Happily, we can more often calculate its derivatives
- Remember Taylor approximations from high school - e.g. 2^{nd} order...

$$f(s_t) \approx f(\bar{s}) + f'(\bar{s})(s_t - \bar{s}) + \frac{1}{2!}f''(\bar{s})(s_t - \bar{s})^2$$

Taylor approximations

In this course, we don't even need second order!

- We will only work with first order approximations
- **In fact, at 1^{st} order we proceed simply by linearizing the equations describing technological constraints and firm/household optimality**
- If we solve those linear equations (like in high school) we will obtain a first order approximation to f
- For more info on 'higher-order asymptotic approximations' and 'perturbation methods' see...
  - https://www.nber.org/econometrics_minicourse_2011/Chapter_2_Pertubation.pdf

So we need to be reminded how to take a linear approximation

Linearization - scalar case

Under various assumptions (that I won't describe here but which hold for the models we consider)

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x})$$

where

$$f'(\bar{x}) \equiv \frac{df}{dx}(\bar{x})$$

This is a first order approximation of f with respect to x, around the point $x = \bar{x}$

Linearization - scalar case

A linear approximation will be exact if f is linear to begin with

- Consider $f(x) = \alpha x$
- $f'(x) = \alpha$ for all x
- $f(x) \overset{Exact}{=} f(\bar{x}) + \alpha(x - \bar{x}) = \alpha x = f(x)$ for all $\bar{x}$
- Clearly, this is pointless

Consider a more general case of a quadratic f

- Consider $f(x) = \frac{\alpha}{2}x^2$
- $f'(x) = \alpha x$
- $f(x) \approx f(\bar{x}) + \alpha\bar{x}(x - \bar{x}) \equiv \hat{f}(x)$ for arbitrary x
- $f(x) \overset{Exact}{=} f(\bar{x}) + \alpha\bar{x}(x - \bar{x}) = f(\bar{x})$ only for $x = \bar{x}$ (trivially)

We take a slope at a point and then using the linear function with *that slope* from *that point*, to approximate the function of interest *at other points*

Linearization - multivariate case

Linear approximation of a scalar valued function with many arguments is essentially the same deal...

- $f(x,y) \approx f(\bar{x},\bar{y}) + \frac{\partial f}{\partial x}(\bar{x},\bar{y})(x-\bar{x}) + \frac{\partial f}{\partial y}(\bar{x},\bar{y})(y-\bar{y}) \equiv \hat{f}(x,y)$
- Consider $f(x,y) = x^2y^3$
- $\frac{\partial f}{\partial x}(\bar{x},\bar{y}) = 2\bar{x}\bar{y}^3$
- $\frac{\partial f}{\partial y}(\bar{x},\bar{y}) = 3\bar{x}^2\bar{y}^2$
- $f(x,y) \approx \bar{x}^2\bar{y}^3 + 2\bar{x}\bar{y}^3(x-\bar{x}) + 3\bar{x}^2\bar{y}^2(y-\bar{y})$

We are making a new function, $\hat{f}$, that will be $= f$ at the approximation point, $(\bar{x},\bar{y})$, but which will only be an approximation for other x and y, by extrapolating the 'slope' of f at $(\bar{x},\bar{y})$.

Linearization - multivariate case

Linear approximation of a scalar valued function with many arguments is essentially the same deal. . .

- $f(x,y) \approx f(\bar{x},\bar{y}) + \frac{\partial f}{\partial x}(\bar{x},\bar{y})(x-\bar{x}) + \frac{\partial f}{\partial y}(\bar{x},\bar{y})(y-\bar{y}) \equiv \hat{f}(x,y)$
- Consider $f(x,y) = x^2y^3$
- $\frac{\partial f}{\partial x}(\bar{x},\bar{y}) = 2\bar{x}\bar{y}^3$
- $\frac{\partial f}{\partial y}(\bar{x},\bar{y}) = 3\bar{x}^2\bar{y}^2$
- $f(x,y) \approx \bar{x}^2\bar{y}^3 + 2\bar{x}\bar{y}^3(x-\bar{x}) + 3\bar{x}^2\bar{y}^2(y-\bar{y})$

We are making a new function, $\hat{f}$, that will be $= f$ at the approximation point, $(\bar{x},\bar{y})$, but which will only be an approximation for other x and y, by extrapolating the 'slope' of f at $(\bar{x},\bar{y})$.

Linearization - multivariate case

Continuing with the $f(x, y) = x^2 y^3$ example...

- Recognize the expression as a function of variables (here x and y)
- Figure out the value each variable takes at the approximation point (here $\bar{x}$ and $\bar{y}$)
- Find the first partial derivative of each function in terms of each variable (here $2xy^3$ and $3x^2y^2$)
    - Note: At this point we have not evaluated those derivatives **at** $\bar{x}$ and $\bar{y}$
- Build the approximation for each function as
    1. The value of the original function at the approximation point
    2. Plus each of the first derivatives **evaluated at the approximation point** $\times$ the deviation of the relevant variable from the approximation point

$$\hat{f}(x, y) \equiv \bar{x}^2\bar{y}^3 + 2\bar{x}\bar{y}^3(x - \bar{x}) + 3\bar{x}^2\bar{y}^2(y - \bar{y})$$

Linearization - multivariate case

Continuing with the $f(x, y) = x^2 y^3$ example...

- Recognize the expression as a function of variables (here x and y)
- Figure out the value each variable takes at the approximation point (here $\bar{x}$ and $\bar{y}$)
- Find the first partial derivative of each function in terms of each variable (here $2xy^3$ and $3x^2y^2$)
  - Note: At this point we have not evaluated those derivatives **at** $\bar{x}$ and $\bar{y}$
- Build the approximation for each function as
  1. The value of the original function at the approximation point
  2. Plus each of the first derivatives **evaluated at the approximation point** $\times$ the deviation of the relevant variable from the approximation point

$$\hat{f}(x, y) \equiv \bar{x}^2\bar{y}^3 + 2\bar{x}\bar{y}^3(x - \bar{x}) + 3\bar{x}^2\bar{y}^2(y - \bar{y})$$

Linearization - multivariate case

Continuing with the $f(x, y) = x^2y^3$ example...

- Recognize the expression as a function of variables (here x and y)
- Figure out the value each variable takes at the approximation point (here $\bar{x}$ and $\bar{y}$)
- Find the first partial derivative of each function in terms of each variable (here $2xy^3$ and $3x^2y^2$)
  - Note: At this point we have not evaluated those derivatives **at** $\bar{x}$ and $\bar{y}$
- Build the approximation for each function as
  1. The value of the original function at the approximation point
  2. Plus each of the first derivatives **evaluated at the approximation point** $\times$ the deviation of the relevant variable from the approximation point

$$\hat{f}(x, y) \equiv \bar{x}^2\bar{y}^3 \; + \; 2\bar{x}\bar{y}^3(x - \bar{x}) + 3\bar{x}^2\bar{y}^2(y - \bar{y})$$

Linearization - common format in our applications

Frequently we will encounter situations where we have (for some functions f and g)

$$f(x,y) = g(x,y)$$

We know that if this relationship is true, then $f(\bar{x},\bar{y}) = g(\bar{x},\bar{y})$ (why?) so our first order linear approximation to this relation

$$\begin{aligned} & f(\bar{x},\bar{y}) + \frac{\partial f}{\partial x}(\bar{x},\bar{y})(x-\bar{x}) + \frac{\partial f}{\partial y}(\bar{x},\bar{y})(y-\bar{y}) \\ \approx\ & g(\bar{x},\bar{y}) + \frac{\partial g}{\partial x}(\bar{x},\bar{y})(x-\bar{x}) + \frac{\partial g}{\partial y}(\bar{x},\bar{y})(y-\bar{y}) \end{aligned}$$

implies

$$\frac{\partial f}{\partial x}(\bar{x},\bar{y})(x-\bar{x}) + \frac{\partial f}{\partial y}(\bar{x},\bar{y})(y-\bar{y}) \approx \frac{\partial g}{\partial x}(\bar{x},\bar{y})(x-\bar{x}) + \frac{\partial g}{\partial y}(\bar{x},\bar{y})(y-\bar{y})$$

Logs and exponentials

A **logarithm** of y 'to base x' is the value to which x must be raised to make it equal y

$$x^{\log_x(y)} \equiv y$$

We will typically be working with the 'natural' logarithm which has the exponential constant, e, as its base

- $e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.71828$
- $\log_e(z)$ is sometimes written $\ln(z)$ but we will typically use $\log(z)$ in this course
- The exponential function $\exp(z) \equiv e^z$
- See https://people.duke.edu/~rnau/411log.htm

Useful properties of the log function

Log of product = sum of logs

$$\log(xy) = \log(x) + \log(y)$$

Exponents become coefficients

$$\log(x^y) = y\log(x)$$

Log of ratio = difference in logs (implied by results above)

$$\log\left(\frac{x}{y}\right) = \log(x) - \log(y)$$

Log of unity = 0 (anything raised to 0 equals unity)

$$\log(1) = 0$$

Useful properties of the log function

Differentiation of logs

$$\frac{d}{dx}\log\left(f(x)\right) = \frac{f'(x)}{f(x)}$$

Differentiation of exponentials

$$\frac{d}{dx}\exp\left(f(x)\right) = f'(x)\exp\left(f(x)\right)$$

Product of exponentials = exponential of sum

$$\exp\left(g(x)\right)\exp\left(f(y)\right) \equiv \exp\left(g(x) + f(y)\right)$$

Useful properties of the log function

$\log(1+i) \approx i$ for small i (useful for gross and net interest rates)

- To see this, take a linear approximation of $\log(1+i)$ around $i = 0$
  - Define $f(i) \equiv \log(1+i)$
  - Then $f(i) \approx \log(1+0) + \frac{1}{1+0}(i-0) = i$
- Note we used $\log(1) = 0$ and $\frac{d}{dx}\log(f(x)) = \frac{f'(x)}{f(x)}$

Difference in logs $\approx$ percentage difference (for small differences)

- To see this note that (recall earlier results)

$$\log(x) - \log(y) = \log\left(\frac{x}{y}\right) = \log\left(1 + \frac{x-y}{y}\right) \approx \frac{x-y}{y}$$

- So today's log minus yesterday's $\approx$ the percentage *growth rate*
- Percent changes are 'unit free' (talk about GDP growth in % not $)

Log linearizations

Log linearization $\Rightarrow$ approximating a function where the slopes and deviations are taken with respect to the *logs of the variables*, rather than the variables themselves

- Often analytically convenient and more natural to think in terms of percent deviations
- For small changes, log deviations are approximately percent deviations

Log linearizations

For a simple way of obtaining a log-linearization I prefer the following:

- Let $x \equiv \log(X)$ and $y \equiv \log(Y)$ (lower case for logs)
- Suppose you're asked to log-linearize $f(X, Y)$ around $(\bar{X}, \bar{Y})$
- Go through $f(X, Y)$ replacing X with $\exp(x)$ and Y with $\exp(y)$
- This effectively defines a new function $\tilde{f}$ such that $\tilde{f}(x, y) \equiv f(X, Y)$
- Then linearize $\tilde{f}$ in terms of x and y around $(\bar{x}, \bar{y})$ where $\bar{x} \equiv \log(\bar{X})$ and $\bar{y} \equiv \log(\bar{Y})$

Log linearizations

Let us go back to our previous example where $f(X, Y) = X^2 Y^3$

- Define a new function in terms of the logs (note the use of $e^{2x} e^{3y} = e^{2x+3y}$)
$$\tilde{f}(x, y) = \exp(2x + 3y)$$
- Linearize $\tilde{f}$ around $(\bar{x}, \bar{y})$

$$\tilde{f}(x, y) \approx \tilde{f}(\bar{x}, \bar{y}) + 2\exp(2\bar{x} + 3\bar{y})(x - \bar{x}) + 3\exp(2\bar{x} + 3\bar{y})(y - \bar{y})$$

We may not be interested in talking in terms of deviations but simply in obtaining expressions in terms of x and y

- Then all the terms involving $\bar{x}$ and $\bar{y}$ will be coefficients on x and y and/or constants

Expectations

We will frequently use the notation $E[\cdot]$ to denote an expectation

- Consider a random variable, X, that follows a discrete distribution
  - $p(i)$ denotes the probability that X takes value $x_i \in \{x_1, \dots x_N\}$

- The expected value (mean) of X is given by the weighted average of its possible values, with the weights given by the probabilities

$$E[X] = \sum_{i=1}^{N} p(i) x_i$$

- The variance is then given by

$$Var(X) = \sum_{i=1}^{N} p(i)(x_i - E[X])^2$$

Expectations

- Consider a random variable, X, that follows a continuous distribution
  - The distribution function, $F(x)$, gives the probability that $X \leq x$
- In this case the expected value is given by the appropriate integral with respect to the distribution function over the support of X, $\mathcal{X}$

$$E[X] = \int_{x \in \mathcal{X}} x dF(x)$$

- The variance is then given by

$$Var(X) = \int_{x \in \mathcal{X}} (x - E[X])^2 dF(x)$$

Properties of the Mean and Variance

For a scalar constants a, b, c

- $E[cX] = cE[X]$
- $E[X + Y] = E[X] + E[Y]$
- $E[XY] = E[XE[Y|X]]$
- $Var(cX) = c^2 Var(X)$
- $Var(X + b) = Var(X)$
- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + abCov(X, Y)$

Normal Distribution

- In this course we will exclusively be dealing with Normal distributions
- If X is Normally distributed with mean μ and variance σ^2 then we write $X \sim N(\mu, \sigma^2)$ and its distribution function is

$$\Phi(x) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

- Perhaps more familiar is the associated probability distribution function

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Normal Distribution

Figure 13: Normal Distribution PDF - Different μ and σ