Macroeconomics II ECON 6140 (Second Half)

Lecture 11
Calibration and Moment Matching

Cornell University
Spring 2025

May 1, 2025

Choosing parameters for a model

Structural macroeconomic models based on maximizing behavior by agents can be used to make predictions about the effects of monetary policy changes, changes in taxes and unemployment benefits etc.

- But model predictions and optimal polices often quantitatively depend on model parameters

How can we choose parameters for the model?

Choosing parameters for a model

Three broad categories of methods:

1. Calibration
2. Moment matching
3. Likelihood based methods

Calibration

Calibration

What does it mean to calibrate a model?

- Using (often micro) evidence from existing studies to choose parameters in preferences, technology and shock specification
- Solve the model and compare model output to actual macro economic data that was not used when choosing the parameters
- If model output closely resembles the actual data, we can claim success

Note: Some use the term *calibration* to simply mean the same thing as parameterizing a model

A “Computational Experiment”

Kydland and Prescott’s definition of a *Computational Experiment:*

1. Ask a well-posed question
2. Write down a suitable model that can address the question
3. Choose parameters from empirical evidence that is not directly related to the question at hand
4. Solve the model
5. Compute the answer

Computational Experiments: Examples

Studies Using Theory	*Question*
Auerbach and Kotlikoff (1987)	What are the effects of the current U.S. social security system on capital formation and intergenerational equity?
Brown, Deardorff and Stern (1994)	What are the potential welfare, wage rate and terms-of-trade effects of NAFTA on Canada, Mexico and the United States?
Gravelle and Kotlikoff (1995)	What are the welfare consequences of the 1986 U.S. tax reform?
Harris (1984)	What are the welfare gains for a small open economy with scale effects and imperfect competition?
Hopenhayn and Rogerson (1993)	What are the welfare costs of a job destruction tax on firms equal to one year's wages?
İmrohoroğlu (1992)	What are the welfare costs of inflation if insurance is imperfect?

Computational Experiments: Examples

Studies Developing Theory	*Question*
Backus, Kehoe and Kydland (1994)	Does theory imply the J-curve pattern of covariance between terms of trade and the trade balance?
Christiano and Eichenbaum (1992b) and Chang (1995)	What is the contribution of public finance shocks to aggregate fluctuations?
Finn (1995)	What is the contribution of oil shocks to business cycle fluctuations?
Greenwood, Hercowitz and Huffman (1988)	Does nonneutrality of technology shocks with respect to the consumption and investment good change the estimate of the contribution of technology shocks to business cycle fluctuations?
Hornstein (1993)	Does the introduction of monopolistic competition into real business cycle models alter the estimate of technology shocks contribution?
Kydland and Prescott (1982)	What is the quantitative nature of fluctuations induced by technology shocks?

Computational Experiments: The Original

Time to build and aggregate fluctuations, by Kydland and Prescott, Econometrica 1982.

Question

- Can an equilibrium model generate realistic business cycles?

Model

- Real Business Cycle model where it takes more than one period to install new capital
- Households choose how much labor to supply and how much to consume
- Investment is income minus consumption

Strategy

- Calibrate model using data
- Solve and simulate model
- Compare model generated data to actual data

Time to build and aggregate fluctuations

For the *calibration stage*, Kydland and Prescott uses data from

- average labor share of GNP to calibrate production function
- investment share of GDP and steady state capital-to-output ratio to set depreciation parameter
- average risk free real interest rate to set the discount rate

Important: Parameters are not chosen based on the features of the data that the model will be used to address

- In addition, some parameters are chosen without direct reference to any empirical observation (e.g. the parameter that governs elasticity of intertemporal substitution).

Time to build: Calibration

TABLE I

MODEL PARAMETERS[a]

Preference Parameters:	$\alpha_0 = 0.50$, $\eta = 0.10$, $\gamma = -0.50$, $\beta = 0.99$
Technology Parameters:	$\nu = 4.0$, $\theta = 0.64$, $\sigma = 0.28 \times 10^{-5}$, $\varphi_1 = \varphi_2 = \varphi_3 = \varphi_4 = 0.25$, $\delta = 0.10$, $\bar{\lambda} = 1.0$
Shock Variances:	$\text{var}(\zeta_1) = 0.0090^2$, $\text{var}(\zeta_2) = 0.0018^2$, $\text{var}(\zeta_3) = 0.0090^2$

[a] For parameters with a time dimension, the unit of time is a quarter of a year.

Time to build: Evaluating the model

TABLE II

AUTOCORRELATIONS OF OUTPUT[a]

Order of Autocorrelations	Model Means (Standard Deviations) of Sample Distribution	U.S. Economy Sample Values for 1950 : 1–1979 : 2
1	.71 (.07)	.84
2	.45 (.12)	.57
3	.28 (.13)	.27
4	.19 (.12)	−.01
5	.02 (.11)	−.20
6	−.13 (.12)	−.30

[a] The length of the sample period both for the model and for the U.S. economy is 118 quarters.

Time to build: Model implied s.d. and correlations

TABLE III

MODEL'S STANDARD DEVIATIONS AND CORRELATIONS WITH REAL OUTPUT[a]

Variable	Standard Deviations: Means (Standard Deviations) of Sample Distribution[b]	Correlations with Output: Means (Standard Deviations) of Sample Distribution
Real Output	1.80 (.23)	—
Consumption	.63 (.09)	.94 (.01)
Investment	6.45 (.62)	.80 (.04)
Inventories	.89 (.06)	−.15 (.11)
Inventories plus	2.00 (.20)	.39 (.06)
Capital Stock	.63 (.08)	−.07 (.06)
Hours	1.05 (.13)	.93 (.01)
Productivity	.90 (.10)	.90 (.02)
Real Interest Rate (Annual)	.23 (.02)	.47 (.10)

[a] The length of the sample period both for the model and for the U.S. economy is 118 quarters.
[b] Measured in per cent.

Time to build: S.d. and correlations in US data

TABLE IV

SAMPLE STANDARD DEVIATIONS AND CORRELATIONS WITH REAL OUTPUT
U.S. ECONOMY 1950 : 1–1979 : 2

	Standard Deviations (per cent)	Correlations with Real Output
Output	1.8	—
Total Consumption	1.3	.74
Services	0.7	.62
Non-Durables	1.2	.71
Durables	5.6	.57
Investment Fixed	5.1	.71
Capital Stock		
Durable Mfg.	1.2	−.21
Non-durable Mfg.	0.7	−.24
Inventories	1.7	.51
Hours	2.0	.85
Productivity	1.0	.10

Calibrating the New Keynesian model

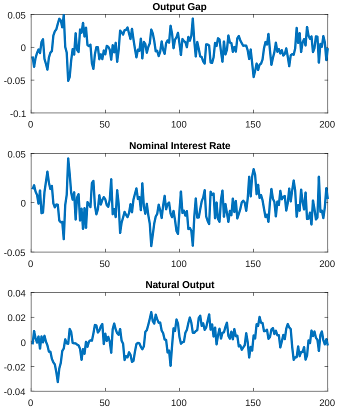
What would a similar exercise look like in the New Keynesian model?

Parameter values:

σ	2	θ	0.75
β	0.99	α	1/3
φ	3	ρ_v	0.5
ε	6	ρ_a	0.75
ϕ_p	1.5	σ_v	0.02
ϕ_y	0.125	σ_a	0.0079

Can we find reasonable values for all parameters?

Simulated data with monetary policy and technology shocks



Data and model moments with technology and monetary policy shocks

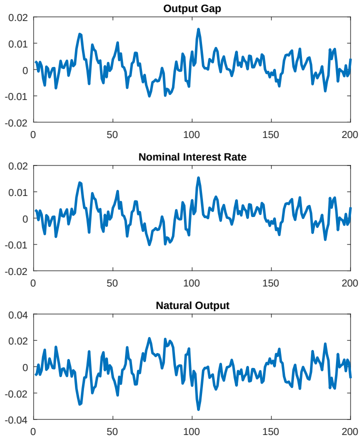
Series	S.d. data (1948-2010)	S.d. model
y_t	0.017	0.016
π_t	0.036	0.005
i_t	0.032	0.014
n_t	0.019	0.01

Shutting down monetary policy shocks

Parameter values:

σ	2	θ	0.75
β	0.99	α	$1/3$
φ	3	ρ_v	0.5
ε	6	ρ_a	0.75
ϕ_p	1.5	σ_v	0
ϕ_y	0.125	σ_a	0.0079

Simulated data with only productivity shocks



Data and model moments with only technology shocks

Series	S.d. data (1948-2010)	S.d. model
y_t	0.017	0.006
π_t	0.036	0.003
i_t	0.032	0.005
n_t	0.019	0.01

Criticism of calibration

- Choice of empirical facts used for calibration is arbitrary
- Choice of which moments that are used to judge the performance of model is arbitrary
- Mapping between micro economic studies to parameters in representative agent model not always straightforward and innocuous
- No formal way to take into account parameter uncertainty
- What is judged a success is not based on any formal criteria

What is the proper role of calibration?

Where there are few theories, or only abstract and unconvincing theories, available and informal exploration in search of new patterns and generalizations is important. A focus on solving and calibrating models, rather than carefully fitting them to data, is reasonable at a stage where solving the models is by itself a major research task. When plausible theories have been advanced, though, and when decisions depend on evaluating them, more systematic collection and comparison of evidence cannot be avoided.

Chris Sims, 1996

Calibration

Summing up

- An relatively informal methodology to choose parameters for a model
- Two stages:
  1. In the calibration stage, some parameters are chosen to match evidence not directly related to main phenomenon studied
  2. In the validation stage, model implied moments are compared to the corresponding data moment
- Calibration has been criticized for allowing for arbitrary judgement calls by researchers (but is sometimes the only feasible approach)

Matching Moments

Matching Moments

What is moment matching?

- Choosing parameters to minimize discrepancy between model and data
- Allows for a more formal statistical treatment of choosing parameters and tests of over-identified restrictions

Note difference from calibration: The data moments are used to choose parameter values, not to evaluate fit of model

Simulated Method of Moments: How to do it

Consider a data moment $h(z_t)$ that has a clear correspondence to a model moment $h\left(y_t, \Theta\right)$. Given simulated model data $\{y_i\}_{i=1}^{N}$ we can compute the discrepancy between the data and model moment as

$$g\left(Z, \Theta\right) = \left(\frac{1}{T}\sum_{t=1}^{T} h(z_t)\right) - \left(\frac{1}{N}\sum_{i=1}^{N} h(y_i, \Theta)\right)$$

We should then choose parameters Θ to minimize the weighted discrepancy between model and data moments

$$\Gamma(\Theta) = g\left(Z, \Theta\right)' \times \left[\Sigma\left(1 + T/N\right)\right]^{-1} \times g\left(Z, \Theta\right)$$

where Σ in the weighting matrix is the covariance of numerical standard deviation of $h(z_t) - h\left(y_t, \Theta\right)$ under the null hypothesis. Since the weighting matrix Σ depends on the parameters Θ, multiple iteration are necessary in order to find the Θ that minimizes $\Gamma(\Theta)$.

Likelihood based estimation

Estimating the parameters in a State Space System

Estimating the parameters in a state space system

For a given state space system

$$\begin{aligned} X_t &= AX_{t-1} + C\mathbf{u}_t : \mathbf{u}_t \sim N(0, I) \quad \text{(state eq.)} \\ Z_t &= DX_t + \mathbf{v}_t : \mathbf{v}_t \sim N(0, \Sigma_{vv}) \quad \text{(measurement eq.)} \end{aligned}$$

How can we find the A, C, D and Σ_v that best fits the data?

The Kalman Filter

For the state space system

$$\begin{array}{rcl} X_t & = & A_t X_{t-1} + C_t \mathbf{u}_t \\ Z_t & = & D_t X_t + \mathbf{v}_t \\ \begin{bmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{bmatrix} & \sim & N\left(\mathbf{0}, \begin{bmatrix} I_n & \mathbf{0}_{n\times l} \\ \mathbf{0}_{l\times n} & \Sigma_{vv} \end{bmatrix}\right) \end{array}$$

the associated Kalman filter is given by

$$X_{t|t} = A_t X_{t-1|t-1} + K_t \left(Z_t - D_t X_{t|t-1}\right)$$

$$K_t = P_{t|t-1} D_t' \left(D_t P_{t|t-1} D_t' + \Sigma_{vv}\right)^{-1}$$

$$\begin{array}{rcl} P_{t+1|t} & = & A_{t+1}\left(P_{t|t-1} - P_{t|t-1} D_{t1}' \left(D_t P_{t|t-1} D_t' + \Sigma_{vv}\right)^{-1} D_t P_{t|t-1}\right) A_{t+1}' \\ & & + C_{t+1} C_{t+1}' \end{array}$$

The innovation sequence can be computed recursively from the innovation representation

$$\widetilde{Z}_t = Z_t - D_t X_{t|t-1}, \quad X_{t+1|t} = A_{t+1} X_{t|t-1} + A_{t+1} K_t \widetilde{Z}_t$$

The likelihood function of a state space model

We can use that the (n-dimensional) vector of innovations $\widetilde{Z}_t$ are conditionally independent Gaussian random vectors to write down the log likelihood function as

$$L(Z \mid \boldsymbol{\Theta}) = (-nT/2)\log(2\pi) - \frac{T}{2}\log|\Omega_t| - \frac{1}{2}\sum_{t=1}^{T}\widetilde{Z}_t'\Omega_t^{-1}\widetilde{Z}_t$$

where

$$\begin{aligned}
\widetilde{Z}_t &= Z_t - DAX_{t-1|t-1} \\
X_{t|t} &= AX_{t-1|t-1} + K_t\left(Z_t - DAX_{t-1|t-1}\right) \\
\Omega_t &= DP_{t|t-1}D' + \Sigma_{vv}
\end{aligned}$$

But how do we find the MLE?

The basic idea

How can we estimate parameters when we cannot maximize likelihood analytically?

We need to

- Be be able to evaluate the likelihood function for a given set of parameters
- Find a way to evaluate a sequence of likelihoods conditional on different parameter vectors so that we can feel confident that we have found the parameter vector that maximizes the likelihood

Numerical maximization of likelihood functions

Numerical maximization

- Grid search
- Steepest ascent
- Newton-Raphson algorithms
- Simulated annealing

Two examples:

- Unobserved components model (Grid search)
- New Keynesian DSGE (Simulated Annealing)

Simple example: Unobserved Components Model

Maximum Likelihood and Unobserved Components Models

Unobserved Component model of inflation

$$
\begin{aligned}
\pi_t &= \tau_t + \eta_t \\
\tau_t &= \tau_{t-1} + \varepsilon_t
\end{aligned}
$$

Decomposes inflation into permanent (τ) and transitory (η) component

- Fits the data well
    - But we may be concerned about having an actual unit root root in inflation on theoretical grounds
- Based on simplified (constant parameters) version of Stock and Watson (JMCB 2007)

The basic formulas

We want to:

1. Estimate the parameters of the system, i.e. estimate σ_η^2 and σ_ε^2
    1.1 Parameter vector is given by $\Theta = \left\{\sigma_\eta^2, \sigma_\varepsilon^2\right\}$
    1.2 $\widehat{\Theta} = \arg\max_{\theta \in \Theta} L(\pi^t \mid \boldsymbol{\Theta})$
2. Find an estimate of the permanent component τ_t at different points in time

The Likelihood function

We have the state space system

$$\begin{aligned}\pi_t &= \tau_t + \eta_t \text{ (measurement equation)}\\ \tau_t &= \tau_{t-1} + \varepsilon_t \text{ (state equation)}\end{aligned}$$

implying that $A = 1, D = 1, C = \sqrt{\sigma_\varepsilon^2}, \Sigma_v = \sigma_\eta^2$. The likelihood function for a state space system is (as always) given by

$$L(Z \mid \mathbf{\Theta}) = -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\Omega_t| - \frac{1}{2}\sum_{t=1}^{T}\widetilde{Z}_t'\Omega_t^{-1}\widetilde{Z}_t$$

where

$$\begin{aligned}\widetilde{Z}_t &= Z_t - DAX_{t-1|t-1}\\ \Omega_t &= DP_{t|t-1}D' + \Sigma_{vv}\end{aligned}$$

and n is the number of observable variables, i.e. the dimension of Z_t.

Starting the Kalman recursions

How can we choose initial values for the Kalman recursions?

- Unconditional variance is infinite because of unit root in permanent component
- A good choice is to choose "neutral" values, i.e. something akin to uninformative priors
  - One such choice is $X_{0|0} = \pi_1$ and $P_{0|0}$ very large (but finite) and constant

$$L(Z \mid \boldsymbol{\Theta}) = -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\Omega_t| - \frac{1}{2}\sum_{t=1}^{T}\widetilde{Z}_t'\Omega_t^{-1}\widetilde{Z}_t$$

Maximizing the Likelihood function

How can we find $\widehat{\Theta} = \arg\max L(\pi^t \mid \mathbf{\Theta})$?

- The dimension of the parameter vector is low so we can use grid search

Define grid for variances σ_ε^2 and σ_η^2

$$\begin{aligned} \sigma_\varepsilon^2 &= \{0, 0.001, 0.002, ..., \sigma_{\varepsilon,\max}^2\} \\ \sigma_\eta^2 &= \{0, 0.001, 0.002, ..., \sigma_{\eta,\max}^2\} \end{aligned}$$

and evaluate the likelihood function for all combinations.

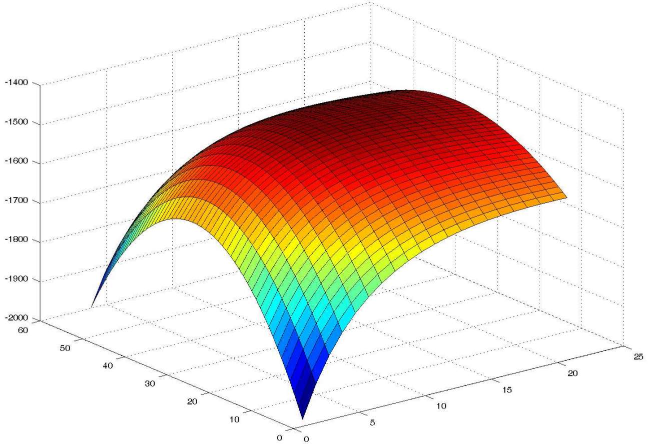
How do we choose boundaries of grid?

- Variances are non-negative
- Both $\widehat{\sigma}_\varepsilon^2$ and $\widehat{\sigma}_\eta^2$ should be smaller than or equal to the sample variance of inflation so we can set $\sigma_{\varepsilon,\max}^2 = \sigma_{\eta,\max}^2 = \frac{1}{T}\sum \pi_t^2$

Grid Search: Fill out the x's

$\sigma_\epsilon^2 \backslash \sigma_\eta^2$	0	0.5	1	1.5	2	2.5
0	x	x	x	x	x	x
0.5	x	x	x	x	x	x
1	x	x	x	x	x	x
1.5	x	x	x	x	x	x
2.5	x	x	x	x	x	x

Maximizing the Likelihood function



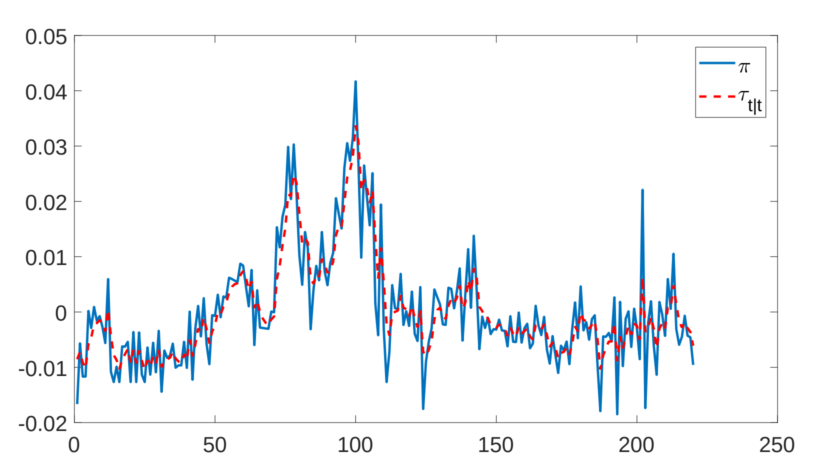
Maximizing the Likelihood function

Estimated parameter values:

- $\widehat{\sigma}_{\varepsilon}^{2} = 0.0028$
- $\widehat{\sigma}_{\eta}^{2} = 0.0051$

We can also estimate the permanent component

Actual Inflation and filtered permanent component



Grid search

Pros:

- With a fine enough grid, grid search always finds the global maximum (if parameter space is bounded)

Cons:

- Computationally infeasible for models with large number of parameters

Steepest ascent based methods

Maximizing the likelihood for larger models

How can we estimate parameters when we cannot maximize likelihood analytically and when grid search is not feasible?

We need to

- Be be able to evaluate the likelihood function for a given set of parameters
- Find a way to evaluate a sequence of likelihoods conditional on difference parameter vectors so that we can feel confident that we have found the parameter vector that maximizes the likelihood

Numerical maximization of likelihood functions

Estimating richer state space models

- Likelihood surface may not be well behaved

We will need more sophisticated maximization routines

Steepest Ascent method

1. Make initial guess of $\Theta = \Theta^{(0)}$
2. Find direction of "steepest ascent" by computing the gradient

$$\mathbf{g}(\Theta) \equiv \frac{\partial \mathcal{L}(Z \mid \Theta)}{\partial \Theta}$$

which is a vector which can be approximated element by element

$$\begin{aligned} & \frac{\partial \mathcal{L}(Z \mid \Theta^{(0)})}{\partial \theta_i} \\ \approx \quad & \frac{\mathcal{L}(Z \mid \theta_j = \theta_j^{(0)} + \varepsilon : j = i; \theta_j = \theta_j^{(0)} \text{ otherwise}) - \mathcal{L}(Z \mid \Theta^{(0)})}{\varepsilon} \end{aligned}$$

for each θ_j in $\Theta = \{\theta_1, \theta_2, ...\theta_J\}$.

3. Take step proportional to gradient, i.e. in the direction of "steepest ascent" by setting new value of parameter vector as $\Theta^{(1)} = \Theta^{(0)} + s\mathbf{g}(\Theta)$

4. Repeat Steps 2 and 3 until convergence.

Steepest Ascent method

Pros:

- Feasible for models with a large number of parameters

Cons:

- Can be hard to calibrate even for simple models to achieve the right rate of convergence
  - Too small steps and “convergence” is achieved to soon
  - Too large step and parameters may be sent off into orbit.
- Can converge on local maximum. (How could a blind man on $K2$ find his way to Mt Everest?)

Newton-Raphson

Newton-Raphson is similar to steepest ascent, but also computes the step size

- Step size depends on second derivative
- May converge faster than steepest ascent
- Requires concavity, so is less robust when shape of likelihood function is unknown

Estimating parameters of the basic New Keynesian model by Maximum Likelihood (Simulated Annealing)

Simulated Annealing, Goffe et al (1994)

- Language is from thermodynamics
- Combines elements of grid search with (strategically chosen) random movements in the parameter space
- Has a good record in practice, but cannot be proven to reach global max quicker than grid search.

Simulated Annealing: The Algorithm

Main inputs: $\Theta^{(0)}$, temperature T, boundaries of Θ, temperature reduction parameter r_T (and the function to be max/minimized $f(\Theta)$).

1. $\theta'_j = \theta_j^{(0)} + r \cdot v_j$ where $r \sim U[-1,1]$ and v_i is an element of the step size vector V.
2. Evaluate $f(\Theta')$ and compare with $f(\Theta^{(0)})$. If $f(\Theta') > f(\Theta^{(0)})$ set $\Theta^{(1)} = \Theta'$. If $f(\Theta') < f(\Theta^{(0)})$ set $\Theta^{(1)} = \Theta'$ with probability $e^{(f(\Theta') - f(\Theta^{(0)})/T}$ and $\Theta^{(1)} = \Theta^{(0)}$ with probability $1-$ $e^{(f(\Theta') - f(\Theta^{(0)})/T}$.
3. After N_s loops through 1 and 2 step length vector V is adjusted in direction so that approx 50% of all moves are accepted.
4. After N_T loops through 1 and 3 temperature is reduced so that $T' = r_T \cdot T$ so that fewer downhill steps are accepted.

Estimating a DSGE model using Simulated Annealing

Remember our benchmark model

$$\begin{aligned}
\pi_t &= \beta E_t \pi_{t+1} + \kappa \widetilde{y}_t \\
\widetilde{y}_t &= E_t \widetilde{y}_{t+1} - \frac{1}{\sigma}\left(i_t - E_t \pi_{t+1} - r_t^n\right) \\
i_t &= \rho + \phi_\pi \pi_t + \phi_y \widehat{y}_t + v_t \\
r_t^n &= \rho - \sigma\left(1 - \rho_a\right)\psi_{ya} a_t
\end{aligned}$$

We want to estimate the parameters in
$\Theta \equiv \{\sigma, \beta, \varphi, \varepsilon, \phi_\pi, \phi_y, \theta, \alpha, \rho_a, \rho_v, \rho_z, \sigma_a, \sigma_v, \sigma_z\}$

Estimating a DSGE model using Simulated Annealing

The solved model can be put in state space form

$$
\begin{aligned}
X_t &= AX_{t-1} + C\mathbf{u}_t \\
Z_t &= DX_t + \mathbf{v}_t
\end{aligned}
$$

where A, C, D and Σ_v are functions of model's "deep" parameters.

The matrices D and the covariance of v_t depends on what observable variables we will include in Z_t.

The log likelihood function of a state space system

For a given state space system

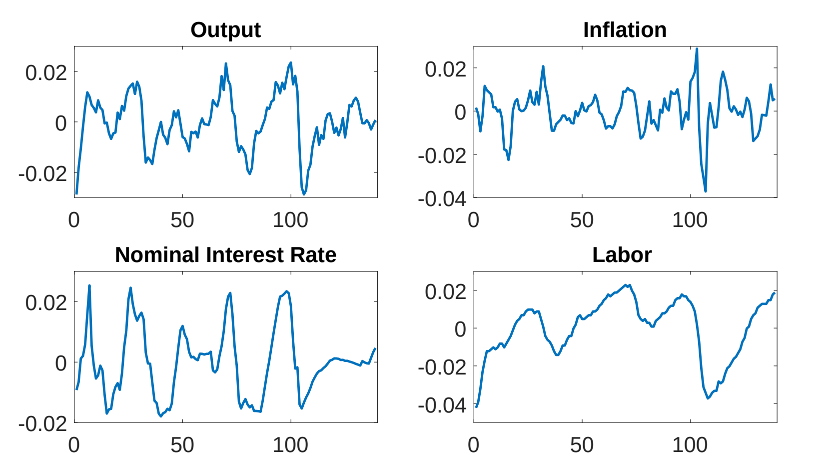
$$\begin{aligned} X_t &= AX_{t-1} + C\mathbf{u}_t \\ Z_t &= DX_t + \mathbf{v}_t \end{aligned}$$

we can evaluate the log likelihood by computing

$$\mathcal{L}(Z \mid \Theta) = -.5 \sum_{t=0}^{T} \left[p \ln(2\pi) + \ln|\Omega_t| + \widetilde{Z}_t' \Omega_t^{-1} \widetilde{Z}_t \right]$$

where $\widetilde{Z}_t$ are the innovation from the Kalman filter

The data: 1983:Q1 - 2017:Q3



Code has three components

1. The main program that defines starting values for simulated annealing algorithm etc
2. A function that translates Θ into a state space system
3. A function that evaluates $\mathcal{L}(Z \mid \Theta)$

Point 2 and 3 are both done by `LLDSGE.m` **Note:** Simulated annealing code in Canvas folder is a *minimizer*

Matlab housekeeping etc

Editor - DSGE_SA.m | Variables - Z

DSGE_SA.m | LLDSGE.m | hpfilter.m | simannb.m | BondAffineMainMCMC.m | filter_gap_DSGE.m

```
% Set up and estimate miniture DSGE model
clc
clear all
close all
global Z
load('Z'); %load data. Order of variables: Inflation, output, interest rate, labor supply

figure(1)
subplot(2,2,1);
plot(Z(1,:),'linewidth',2);title('Output','fontsize',16);
subplot(2,2,2);
plot(Z(2,:),'linewidth',2);title('Inflation','fontsize',16)
subplot(2,2,3);
plot(Z(3,:),'linewidth',2);title('Nominal Interest Rate','fontsize',16)
subplot(2,2,4);
plot(Z(4,:),'linewidth',2);title('Labor ','fontsize',16)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Define starting values and boundaries for Θ

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Initial values of structural parameters
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
tau       = 3;  % Since I am using sigma for the standard deviaion of the shocks, I am using t
beta      = 0.99; % discount factor
theta     = 3/4; % degree of price stickiness
phi_pi    = 1.5; % taylor rule parameter
phi_y     = 0.125; % taylor rule parameter
varphi    = 3; %inverse of elastiicity of labor supply
alpha     = 1/3; %production function parameter
eps       = 6; % elasticity of substitution between goods i and j in the consumption basket
rho_v     = 0.5; %persistence parameter
rho_a     = 0.75; %persistence parameter
sigma_v   = 0.02; %standard deviation
sigma_a   = (0.012^2*(1-rho_a^2))^.5; %standard deviation of innovation to a_t


THETA=[tau,beta,theta,phi_pi,phi_y,varphi,alpha,eps,rho_v,rho_a,sigma_v,sigma_a;]';%Starting v
LB=[0,0,0,1,0,1,0,1,zeros(1,4)]';%Lower bound for parameter vector
UB=[10,1,1,5,5,10,1,25,1,1,10,10]';%Upper bound for parameter vector
x=THETA;
```

Start the simulated annealing algorithm and print output

```
% x=THETA;
sa_t= 5;
sa_rt=.3;
sa_nt=5;
sa_ns=5;
% warning off all;

[xhat]=simannb( 'LLDSGE', x, LB, UB, sa_t, sa_rt, sa_nt, sa_ns, 1);

%-----------------------------------------------------------------------
thetalabel=['tau    ';'beta   ';'theta  ';'phi_pi ';'phi_y  ';'varphi ';'alpha  ';'eps    ';'r
disp('ML estimate of THETA')
disp([thetalabel, num2str(xhat)])
%-----------------------------------------------------------------------
a_hat = filter_gap_DSGE(xhat,Z);
figure(2)
plot(Z(1,:),'linewidth',2);hold on;plot(a_hat,'linewidth',2)
legend('Output','Output gap')
```

Running simulated annealing I

C: ▸ Users ▸ pkn8 ▸ Dropbox ▸ Teaching ▸ Graduate ▸ 1st Year Macro ▸ Ma

Current Folder | Command Window

```
initial loss function value:
  -1.9568e+03

No. of evaluations
    61

  current temperature
     5

current optimum function value
  -2.1990e+03

No. of downhill steps
     9

No. of accepted uphill steps
     2

No. of rejections
    49

current optimum vector
    3.0000
    0.5580
    0.5898
    3.1080
    0.1250
    5.9375
    0.1714
    6.6584
    0.5000
    0.7500
```

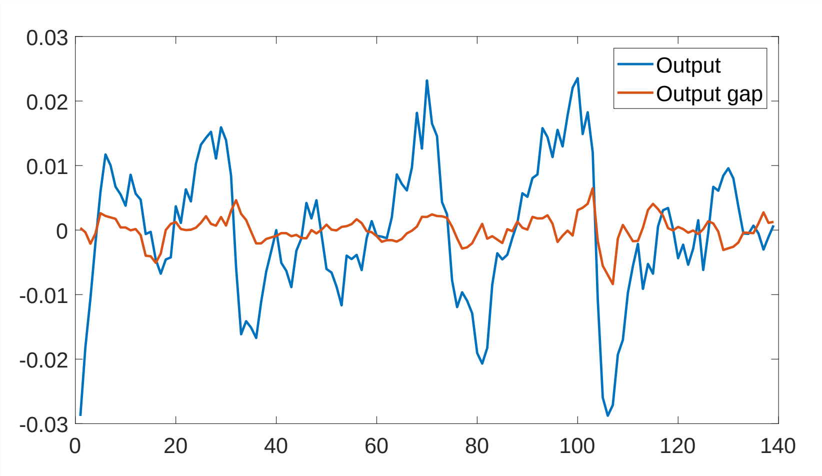
Running simulated annealing II

```
Elapsed time is 13.834891 seconds.
No. of evaluations
        2041

  current temperature
    0.0036

current optimum function value
  -2.2332e+03

No. of downhill steps
     6

No. of accepted uphill steps
     8

No. of rejections
   226

current optimum vector
    3.2877
    0.7271
    0.2900
    3.4086
    1.6229
    4.7205
    0.4325
   19.3046
    0.7482
    0.8662
    0.0218
    0.0059

Elapsed time is 14.216949 seconds.
No. of evaluations
```

Running simulated annealing III

```
Elapsed time is 20.444027 seconds.
simulated annealing achieved termination after 3001 evals
optimum function value
  -2.2332e+03

ML estimate of THETA
tau         3.2877
beta      0.727051
theta     0.290015
phi_pi     3.40862
phi_y      1.62294
varphi     4.72045
alpha       0.4325
eps        19.3046
rho_v     0.748235
rho_a      0.86618
sigmav  0.0217664
sigmaa 0.00589684
>>
```

Output and the estimated output gap at MLE



Summing up

DSGE model

- Input: Vector of parameters Θ
- Output: A state space system

$$f(\Theta) \to \{A, C, D, \Sigma_{vv}\}.$$

Kalman filter

- Estimate latent variables in state space system
- Evaluate the likelihood function for given parameterized state space system

Optimization algorithm

- Find parameter vector Θ that maximizes likelihood given the data

References

Essential

- Kydland, F.E. and Prescott, E.C., 1996. The computational experiment: an econometric tool. The Journal of Economic Perspectives, 10(1), pp.69-85.
- Sims, C.A., 1996. Macroeconomics and methodology. The Journal of Economic Perspectives, 10(1), pp.105-120.

Extra

- Interview with Thomas J Sargent in Macroeconomic Dynamics
- DeJong, D.N. and Dave, C., 2011. Structural macroeconometrics. Princeton University Press.
- Kydland, F.E. and Prescott, E.C., 1982. Time to build and aggregate fluctuations. Econometrica, pp.1345-1370.
- Hansen, L.P. and Heckman, J.J., 1996. The empirical foundations of calibration. The Journal of Economic Perspectives, 10(1), pp.87-104.
- Chapter 5 in Hamilton, J.D., 1994. *Time series analysis*. Princeton: Princeton university press.
- Goffe, W.L., Ferrier, G.D. and Rogers, J., 1994. *Global optimization of statistical functions with simulated annealing*. Journal of econometrics, 60(1-2), pp.65-99.