# Dynamic Choice Models with Conditional Gauss Markov Signals – Theory and Perturbation Approximation

Ivana Komunjer, Elena Afanasyeva and Molin Zhong

Georgetown University, Federal Reserve Board and Federal Reserve Board

# Motivation

Empirical macroeconomics: there is a widely shared view that (1) time-varying economic uncertainty and (2) agents' beliefs (expectations) play an important role in driving macroeconomic fluctuations.

Theoretical question: how to solve dynamic choice (DSGE) models in which agents have incomplete information about the model states and thus form beliefs about those states?

Computational question: Is the perturbation method still applicable when (some of) the model states remain hidden to the decision maker?

# Literature

- Macroeconomics Models with Incomplete Information:
  - LQ approximation: Kydland and Prescott (1982)
  - log-linearization: Barsky and Sims (2012), Blanchard, L'Huillier, and Lorenzoni (2013)
- Perturbation Method:
  - Optimal control: Fleming (1971), Bensoussan (1988)
  - Economics: Judd (1996, 1998)
- Partially Observed Stochastic Optimal Control:
  - Streibel (1975), Bertsekas (1976), Bertsekas and Shreve (1996)
- Small Noise Limits:
  - continuous time: Hijab (1984), James (1991), Baras, Bensoussan and James (1998)
  - discrete time: James, Baras and Elliott (1994)

# This Talk:

Consider a simple RBC model with information structure a la Cogley and Sargent (2005), Primiceri (2005).

This is a special case of a large class of nonlinear models which we call conditional Gauss Markov.

Using existing control theory results, we calculate the full information equivalent of the model. In the full information equivalent all the states are observed by the agent.

We then derive the limit of this model as the noise goes to zero and show that the perturbation method applies.

1. Motivation
2. Sequential Choice Setup
3. Conditional Gauss Markov Process
4. Full Information Equivalence
5. Perturbation Approximation
6. Conclusion

# Stochastic Growth Model with Incomplete Information

The agent's problem is to choose consumption $C_t$ and savings $K_t$ to maximize

$$E\left[\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\nu}}{1-\nu}\right], \quad 0<\beta<1 \tag{1}$$

subject to the output $Y_t$ production constraint

$$\begin{aligned} Y_t &= A_t K_{t-1}^{\alpha} && (2)\\ Y_t &= C_t + K_t - (1-\delta)K_{t-1} && (3)\end{aligned}$$

with $0<\delta<1$ and $0<\alpha<1$.

# Stochastic Growth Model with Incomplete Information

Similar to Cogley and Sargent (2005), Primiceri (2005): $a_t = \ln A_t$ evolves according to the process:

$$\begin{aligned} a_{t+1} &= \theta_t a_t + \sigma_{a,t+1}\epsilon_{a,t+1}, \quad \epsilon_{a,t+1} \sim \text{ iid } N(0,1) \\ \theta_t &= \theta_{t-1} + \sigma_\theta \epsilon_{\theta t}, \qquad \epsilon_{\theta t} \sim \text{ iid } N(0,1) \end{aligned}$$

where

$$\ln \sigma^2_{a,t+1} = \omega_a + \lambda_a \ln \sigma^2_{at} + \alpha_a a_t, \quad \sigma^2_{a0} \text{ given},$$

At each $t \geq 0$, the agent observes $a_t$ but not $\theta_t$. We assume that $\{\epsilon_{at}\}_{t=1}^{\infty}$, $\{\epsilon_{\theta t}\}_{t=1}^{\infty}$, $a_0$, and $\theta_0 \sim N(\overline{\theta}, \overline{P})$ are independent.

Interpretation: there is both a drifting unobserved parameter $\theta_t$ and a time varying uncertainty in productivity $\sigma_{at}$

# Stochastic Growth Model with Incomplete Information

First question: how should one solve for the optimal consumption $C_t$?

Second question: is it possible to approximate the optimal policy using the *perturbation* method?

# Stochastic Growth Model with Incomplete Information

Why is this problem difficult?

The states are:

$$\underbrace{K_{t-1}, a_t,}_{\text{observed}} \underbrace{\theta_t}_{\text{hidden}}$$

- The decision maker has incomplete information about the model states and thus forms beliefs about those states
  - LQ models (Kushner, 1971): numerous applications in macroeconomics (Kydland and Prescott, 1982, Barsky and Sims, 2012, Blanchard et al., 2013)
  - nonlinear models: the state becomes infinite dimensional (posterior distribution of the hidden state given the observables)

# Stochastic Growth Model with Incomplete Information

Why is this problem difficult?

The states are:

$$\underbrace{K_{t-1}, a_t,}_{\text{observed}} \underbrace{\theta_t}_{\text{hidden}}$$

- Existing nonlinear solution methods require the model states to be observed by the decision maker:
  - discrete approximation methods: discretization methods that aim to approximate the Bellman operator (e.g. Tauchen and Hussey, 1991, Rust, 1996, 1997)
  - smooth approximation methods (e.g. Judd, 1996): projection (global), perturbation (local)

# Stochastic Growth Model with Incomplete Information

- First contribution: we derive full information equivalents of incomplete information dynamic choice models with conditional Gauss Markov information structure
  - model state is finite dimensional even though the model is nonlinear
- Second contribution: we derive the small noise limit of the model which is needed to apply the *perturbation* method around the deterministic model solution

1 Motivation

2 Sequential Choice Setup

3 Conditional Gauss Markov Process

4 Full Information Equivalence

5 Perturbation Approximation

6 Conclusion

# Conditional Gauss Markov Process

Recall first the well-known linear Gaussian state space model:

$$\xi_{t+1} = A\xi_t + u + Bw_{t+1} \quad \text{(transition)}$$
$$z_t = C\xi_t + v + Dw_t \quad \text{(observation)}$$

with

- $w_t \sim \ iid\ N(0, Id)$, $\xi_0 \sim N(\bar{\bar{\xi}}_0, \bar{\Sigma}_0)$, $\xi_0 \perp \{w_t\}_{t=0}^{\infty}$
- $z_t$ (dimension $n_z$) is observable.
- $\xi_t$ and $w_t$ (dimensions $n_\xi$ and $n_w$) are latent.
- $A, B, C, D, u, v$ are conformable system matrices.

# Conditional Gauss Markov Process

Recall first the well-known linear Gaussian state space model:

$$\begin{aligned} \xi_{t+1} &= A\xi_t + u + Bw_{t+1} && \text{(transition)} \\ z_t &= C\xi_t + v + Dw_t && \text{(observation)} \end{aligned}$$

Key properties:

- let $\xi^t \equiv \{\xi_t, \xi_{t-1}, \ldots, \xi_0\}$, $z^t \equiv \{z_t, z_{t-1}, z_{t-2}, \ldots, z_0\}$; then

$$(\xi^t, z^t) \quad \text{is (multivariate) Gaussian for } t = 0, 1, \ldots$$

- in particular, this implies that for all $t = 0, 1, \ldots$
  - $\xi_t$ and $z_t$ are Gaussian (marginals)
  - $\xi_t \mid z^t$ and $z_{t+1} \mid z^t$ are Gaussian (conditionals)

# Conditional Gauss Markov Process

Recall first the well-known linear Gaussian state space model:

$$\begin{aligned} \xi_{t+1} &= A\xi_t + u + Bw_{t+1} && \text{(transition)} \\ z_t &= C\xi_t + v + Dw_t && \text{(observation)} \end{aligned}$$

1. Conditional Gaussian $\xi_t \mid z^t$:
   - sufficient statistics:

$$\hat{\xi}_t \equiv E(\xi_t \mid z^t) \quad \text{and} \quad \Sigma_t \equiv E[(\xi_t - \hat{\xi}_t)(\xi_t - \hat{\xi}_t)' \mid z^t]$$

   - well-known Kalman filtering equations give:

$$(\hat{\xi}_0, \Sigma_0) \xrightarrow{z_1} (\hat{\xi}_1, \Sigma_1) \xrightarrow{z_2} (\hat{\xi}_2, \Sigma_2) \xrightarrow{z_3} \cdots$$

# Conditional Gauss Markov Process

Recall first the well-known linear Gaussian state space model:

$$\begin{aligned} \xi_{t+1} &= A\xi_t + u + Bw_{t+1} && \text{(transition)} \\ z_t &= C\xi_t + v + Dw_t && \text{(observation)} \end{aligned}$$

2. Markovian sufficient statistics:
   - $(\hat{\xi}_t, \Sigma_t)$ is a function of $(\hat{\xi}_{t-1}, \Sigma_{t-1})$ and $z_t$
   - the sequence $\{(\hat{\xi}_t, \Sigma_t)\}_{t=0}^{\infty}$ is Markovian or transitive (Bahadur, 1954; Shiryaev, 1964, 1969)
   - key property: optimal controls can be obtained as functions of these statistics (optimal strategies are *memoryless* in the sense of Blackwell, 1964)

# Conditional Gauss Markov Process

Recall first the well-known linear Gaussian state space model:

$$\begin{aligned} \xi_{t+1} &= A\xi_t + u + Bw_{t+1} && \text{(transition)} \\ z_t &= C\xi_t + v + Dw_t && \text{(observation)} \end{aligned}$$

Question: are there more general models in which the statistics $(\hat{\xi}_t, \Sigma_t)$ are Markovian/transitive sufficient statistics?

Answer: YES!

# Conditional Gauss Markov Process

Define a nonlinear state space model:

$$\begin{aligned} \xi_{t+1} &= A(z_t, m_t)\xi_t + u(z_t, m_t) + B(z_t, m_t)w_{t+1} \quad \text{(transition)} \\ z_t &= C(z_{t-1}, m_{t-1})\xi_t + v(z_{t-1}, m_{t-1}) + D(z_{t-1}, m_{t-1})w_t \quad \text{(observatio} \end{aligned}$$

with

$$m_t = f(m_{t-1}, z_{t-1})$$

where

- the entries of $A(z_t, m_t), B(z_t, m_t), C(z_t, m_t), D(z_t, m_t)$ and $u(z_t, m_t), v(z_t, m_t)$ are finite with probability one
- $m_t$ is observed (dimension $n_m \geq 0$)
- $m_0$ given (deterministic)

# Conditional Gauss Markov Process

Define a nonlinear state space model:

$$\begin{aligned} \xi_{t+1} &= A(z_t, m_t)\xi_t + u(z_t, m_t) + B(z_t, m_t) w_{t+1} \quad \text{(transition)} \\ z_t &= C(z_{t-1}, m_{t-1})\xi_t + v(z_{t-1}, m_{t-1}) + D(z_{t-1}, m_{t-1}) w_t \quad \text{(observatior} \end{aligned}$$

with

$$m_t = f(m_{t-1}, z_{t-1})$$

1. Conditional Gaussian $\xi_t \mid z^t$:
   - $\xi_t \mid z^t$ and $z_{t+1} \mid z^t$ are still conditional Gaussian
   - however, $(\xi^t, z^t)$ is no longer (multivariate) Gaussian; thus the marginals of $\xi_t$ and $z_t$ are no longer Gaussian

# Conditional Gauss Markov Process

Define a nonlinear state space model:

$$
\begin{aligned}
\xi_{t+1} &= A(z_t, m_t)\xi_t + u(z_t, m_t) + B(z_t, m_t)w_{t+1} \quad \text{(transition)} \\
z_t &= C(z_{t-1}, m_{t-1})\xi_t + v(z_{t-1}, m_{t-1}) + D(z_{t-1}, m_{t-1})w_t \quad \text{(observatio}
\end{aligned}
$$

with

$$
m_t = f(m_{t-1}, z_{t-1})
$$

2. Markovian sufficient statistics:
   - $(\hat{\xi}_0, \Sigma_0, m_0) \xrightarrow{z_1} (\hat{\xi}_1, \Sigma_1, m_1) \xrightarrow{z_2} (\hat{\xi}_2, \Sigma_2, m_2) \xrightarrow{z_3} \cdots$
   - Kalman filtering equations still hold but the system matrices are now random

# Conditional Gauss Markov Process

Define a nonlinear state space model:

$$\begin{aligned} \xi_{t+1} &= A(z_t, m_t)\xi_t + u(z_t, m_t) + B(z_t, m_t)w_{t+1} \quad \text{(transition)} \\ z_t &= C(z_{t-1}, m_{t-1})\xi_t + v(z_{t-1}, m_{t-1}) + D(z_{t-1}, m_{t-1})w_t \quad \text{(observatio} \end{aligned}$$

with

$$m_t = f(m_{t-1}, z_{t-1})$$

We call this a Conditional Gauss Markov model. The resulting process is conditionally Gaussian (Liptser and Shiryaev, 1978) and its sufficient statistics are Markovian/transitive (our contribution).

# Examples

1. Time-varying volatility: GARCH models

$$z_t = \xi_t + \underbrace{\sigma_{zt}}_{D(z_{t-1}, m_{t-1})} \epsilon_{z,t}$$

$$\xi_{t+1} = \underbrace{\rho}_{A} \xi_t + \underbrace{\sigma_\xi}_{B} \epsilon_{\xi,t}$$

$$\sigma_{zt}^2 = \omega + \beta \sigma_{zt-1}^2 + \alpha z_{t-1}^2$$

here

- $m_t = \sigma_{zt}$
- $f(m_{t-1}, z_{t-1}) = [\omega + \beta m_{t-1}^2 + \alpha z_{t-1}^2]^{1/2}$

# Examples

2. Regime switching

$$z_t = \xi_t + \underbrace{\sigma_{z,t-1}}_{D(z_{t-1,m_{t-1}})} \epsilon_{z,t}$$

$$\xi_{t+1} = \underbrace{\rho_t}_{A(z_t,m_t)} \xi_t + \underbrace{\sigma_{\xi t}}_{B(z_t,m_t)} \epsilon_{\xi,t}$$

$$\{\sigma_{\xi t}, \rho_t, \sigma_{zt}\} = \begin{cases} \{\sigma_{\xi 1}, \rho_1, \sigma_{z1}\} \text{ for } \Delta z_t > \bar{z} \\ \{\sigma_{\xi 2}, \rho_2, \sigma_{z2}\} \text{ for } \Delta z_t \leq \bar{z} \end{cases}$$

here

- $m_t = z_{t-1}$

# Examples

3. Exogenously Evolving Time-Varying Parameters

$$z_t = (1-\xi_t)z^* + \xi_t z_{t-1} + \sigma_z \varepsilon_{zt}$$
$$\xi_t = (1-\lambda)\xi^* + \lambda \xi_{t-1} + \sigma_\xi \varepsilon_{\xi t}$$

Re-write as:

$$z_t = \underbrace{(z_{t-1} - z^*)}_{C(z_{t-1})} \xi_t + \underbrace{z^*}_{v} + \underbrace{\sigma_z}_{D} \varepsilon_{zt}$$
$$\xi_t = \underbrace{\lambda}_{A} \xi_{t-1} + \underbrace{(1-\lambda)\xi^*}_{u} + \underbrace{\sigma_\xi}_{B} \varepsilon_{\xi t}$$

Here $m_t = \emptyset$.

1 Motivation

2 Sequential Choice Setup

3 Conditional Gauss Markov Process

4 Full Information Equivalence

5 Perturbation Approximation

6 Conclusion

# Back to the Sequential Choice Model

$$\max E\left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\nu}}{1-\nu}\right], \quad 0 < \beta < 1$$

subject to

$$\begin{aligned} C_t + K_t - (1-\delta)K_{t-1} &= A_t K_{t-1}^{\alpha} \\ a_{t+1} &= \theta_t a_t + \sigma_{a,t+1}\epsilon_{a,t+1} \\ \theta_t &= \theta_{t-1} + \sigma_\theta \epsilon_{\theta t} \\ \ln \sigma_{a,t+1}^2 &= \omega_a + \lambda_a \ln \sigma_{at}^2 + \alpha_a a_t \end{aligned}$$

# Control and Filtering Densities

Euler equation:

$$1 = E\left[\left(\frac{C(K_t, a_{t+1}, p_{\theta t+1})}{C_t}\right)^{-\nu} \beta\left((1-\delta) + \alpha \exp(a_{t+1}) K_t^{\alpha-1}\right) \middle| I_t\right]$$

$p_{\theta t+1}$ is the conditional density of $\theta_{t+1}$ given $I_{t+1}$ where $I_{t+1} = I_t \cup \{Y_{t+1}, a_{t+1}, C_t, K_t\}$. We refer to this density as the filtering density.

Thus, the term inside the conditional expectation is a function of $a_{t+1}$ and $I_t$ and the expectation is with respect to the conditional density $p_{a_{t+1,t}}$ of $a_{t+1}$ given $I_t$. We refer to this density as the control density (for the terminology, see Streibel, 1975).

# Full Information Equivalence

In Conditional Gauss Markov models, both the filtering and control densities are conditionally Gaussian with statistics $(\hat{\xi}_t, \Sigma_t, m_t)$ that are Markovian/transitive

Here, $p_{\theta t}$ is a function of

$$\hat{\xi}_t = E(\theta_t \mid a^t) = \hat{\theta}_t$$
$$\Sigma_t = E[(\theta_t - \hat{\theta}_t)^2 \mid a^t] = P_t$$
$$m_t = \sigma^2_{at}$$

and $p_{a_{t+1,t}}$ can be calculated from the *innovations* representation of the conditional Gauss Markov process

# Full Information Equivalence

In the full information equivalent model, we are looking for an optimal policy $C(K_{t-1}, a_t, \hat{\theta}_t, P_t, \sigma^2_{at})$ which solves the Euler equation:

$$E\left[\beta\left(\alpha \exp(a_{t+1}) K_t^{\alpha-1} + (1-\delta)\right)\left(\frac{C_{t+1}}{C_t}\right)^{-\nu} \middle| K_{t-1}, a_t, \hat{\theta}_t, P_t, \sigma^2_{at}\right] = 1$$

# Full Information Equivalence

The full information equivalent states $(K_{t-1}, a_t, \hat{\theta}_t, P_t, \sigma^2_{at})$ are all observed by the agent and evolve as:

$$K_t = (1-\delta)K_{t-1} + (\exp a_t)K^{\alpha}_{t-1} - C(K_{t-1}, a_t, \hat{\theta}_t, P_t, \sigma^2_{at})$$

$$a_{t+1} = \hat{\theta}_t a_t + \sqrt{a_t^2 P_t + \sigma^2_{a,t+1}}\epsilon_{t+1}$$

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{a_t P_t}{\sqrt{a_t^2 P_t + \sigma^2_{a,t+1}}}\epsilon_{t+1}$$

$$P_{t+1} = \sigma^2_{\theta} + \frac{\sigma^2_{a,t+1}}{a_t^2 P_t + \sigma^2_{a,t+1}} P_t$$

$$\sigma^2_{a,t+1} = \exp(\omega_a + \lambda_a \ln \sigma^2_{at} + \alpha_a a_t)$$

with $\{\epsilon_t\}_{t=1}^{+\infty} \sim$ iid $N(0,1)$

# Full Information Equivalence

Interpretation: in incomplete information models with conditional Gauss Markov information structure, the model states are

- conditional expectations $\hat{\theta}_t$ (forecastable component)
- conditional variance $P_t$ (Jurado, Ludvingson, Ng, 2015, measure of uncertainty)
- other time-varying variables (here time-varying volatility)

1 Motivation

2 Sequential Choice Setup

3 Conditional Gauss Markov Process

4 Full Information Equivalence

5 Perturbation Approximation

6 Conclusion

# Perturbation Approximation

Assume now that instead of being standardized, the shocks now follow

$$\begin{aligned} a_{t+1} &= \theta_t a_t + \sigma_{a,t+1}\epsilon_{a,t+1}, \quad \epsilon_{a,t+1} \sim \text{ iid } N(0,\sigma^2) \\ \theta_t &= \theta_{t-1} + \sigma_\theta \epsilon_{\theta t}, \qquad \epsilon_{\theta t} \sim \text{ iid } N(0,\sigma^2) \end{aligned}$$

where

$$\ln \sigma^2_{a,t+1} = \omega_a + \lambda_a \ln \sigma^2_{at} + \alpha_a a_t, \quad \sigma^2_{a0} \text{ given}$$

Question: how does the full information equivalent of this model depend on $\sigma$?

# Perturbation Approximation

Let

$$\Pi_t \equiv \frac{E[(\theta_t - \hat{\theta}_t)^2 \mid a^t]}{\sigma^2}$$

then we have the following recursion for the scaled conditional variances

$$\Pi_{t+1} = \sigma_\theta^2 + \frac{\sigma_{a,t+1}^2}{a_t^2 \Pi_t + \sigma_{a,t+1}^2} \Pi_t$$

which does not depend on $\sigma^2$.

# Perturbation Approximation

The exogenous states of the full information equivalent model can then be written as:

$$a_{t+1} = \hat{\theta}_t a_t + \sqrt{a_t^2 \Pi_t + \sigma_{a,t+1}^2}\epsilon_{t+1}$$

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{a_t \Pi_t}{\sqrt{a_t^2 \Pi_t + \sigma_{a,t+1}^2}}\epsilon_{t+1}$$

$$\Pi_{t+1} = \sigma_\theta^2 + \frac{\sigma_{a,t+1}^2}{a_t^2 \Pi_t + \sigma_{a,t+1}^2}\Pi_t$$

$$\sigma_{a,t+1}^2 = \exp(\omega_a + \lambda_a \ln \sigma_{at}^2 + \alpha_a a_t)$$

with $\{\epsilon_t\}_{t=1}^{+\infty} \sim$ iid $N(0, \sigma^2)$

This form is precisely the starting point of the perturbation approximation.

# Perturbation Approximation

This results hold for general conditional Gauss Markov models:

- scaled conditional variances $\sigma^{-1}\Sigma_t$ follow a recursion which does not depend on $\sigma$
- conditional means $\hat{\theta}_t$ are invariant to $\sigma$

An immediate implication of this result is that the deterministic limit ($\sigma \to 0$) of the incomplete information model is *not* the same as that of the complete information model in which the decision maker observes $\theta_t$.

# Conclusion

We consider the problem of optimal choice in a dynamic stochastic model with incomplete information.

If the information structure in the model is conditional Gauss Markov, then the original model has a finite dimensional full information equivalent, whose solution can be approximated using the perturbation method.

What is the practical relevance of our results? We provide a method by which nonlinear solutions to DSGE models with incomplete information can be obtained.