Makroekonomia Zaawansowana

Ćwiczenia 9 – Estymacja modeli DSGE za pomocą metod bayesowskich

Andrzej Torój

SGH Szkoła Główna Handlowa w Warszawie

Instytut Ekonometrii |
Zakład Ekonometrii Stosowanej

Plan ćwiczeń

Plan prezentacji

Ekonometria bayesowska – naczelna idea

- w klasycznej ekonometrii:
  - cała wiedza o wartości parametru pochodzi z próby statystycznej
  - szacujemy parametr punktowo, ew. z przedziałem ufności
- w bayesowskiej ekonometrii:
  - parametr traktujemy nie jako jedną "prawdziwą" wartość, ale jako zmienną losową (tzn. może przyjmować wiele różnych wartości z określonym prawdopodobieństwem – otrzymujemy ROZKŁAD)
  - wiedza o wartości parametru pochodzi częściowo z próby, ale częściowo jest narzucona z góry (tzw. rozkład *a priori*)

Dlaczego metody bayesowskie w DSGE?

Rozwiązany zlinearyzowany model DSGE jest przecież postaci $y_t = \mathrm{M}y_{t-1} + \mathrm{N}f_t$ i można go szacować metodą największej wiarygodności tak jak model SVAR...

- silne restrykcje teoretyczne dla ekonomicznie interpretowalnych wartości parametrów
- ucieczka parametrów w bezsensowne ekonomicznie zakresy grozi złamaniem warunków Blancharda-Kahna
- w badaniach empirycznych często dość krótkie próby – potrzeba informacji *a priori*
- model DSGE jest na ogół zbyt silnie wyidealizowany (zawiera zbyt mało sprzężeń w porównaniu z rzeczywistością), aby bez "wsparcia" *a priori* być skonfrontowanym z danymi (np. w ramach metody FIML)
- wiedza spoza próby statystycznej jest wprowadzana w sposób systematyczny, a nie nieformalny
- czasami mamy w modelu więcej zmiennych obserwowalnych niż wstrząsów – przy metodach klasycznych to problem

Dlaczego metody bayesowskie w DSGE?

Rozwiązany zlinearyzowany model DSGE jest przecież postaci $y_t = \mathrm{M}y_{t-1} + \mathrm{N}f_t$ i można go szacować metodą największej wiarygodności tak jak model SVAR...

- silne restrykcje teoretyczne dla ekonomicznie interpretowalnych wartości parametrów
- ucieczka parametrów w bezsensowne ekonomicznie zakresy grozi złamaniem warunków Blancharda-Kahna
- w badaniach empirycznych często dość krótkie próby – potrzeba informacji *a priori*
- model DSGE jest na ogół zbyt silnie wyidealizowany (zawiera zbyt mało sprzężeń w porównaniu z rzeczywistością), aby bez "wsparcia" *a priori* być skonfrontowanym z danymi (np. w ramach metody FIML)
- wiedza spoza próby statystycznej jest wprowadzana w sposób systematyczny, a nie nieformalny
- czasami mamy w modelu więcej zmiennych obserwowalnych niż wstrząsów – przy metodach klasycznych to problem

Dlaczego metody bayesowskie w DSGE?

Rozwiązany zlinearyzowany model DSGE jest przecież postaci $y_t = \mathrm{M}y_{t-1} + \mathrm{N}f_t$ i można go szacować metodą największej wiarygodności tak jak model SVAR...

- silne restrykcje teoretyczne dla ekonomicznie interpretowalnych wartości parametrów
- ucieczka parametrów w bezsensowne ekonomicznie zakresy grozi złamaniem warunków Blancharda-Kahna
- w badaniach empirycznych często dość krótkie próby – potrzeba informacji *a priori*
- model DSGE jest na ogół zbyt silnie wyidealizowany (zawiera zbyt mało sprzężeń w porównaniu z rzeczywistością), aby bez "wsparcia" *a priori* być skonfrontowanym z danymi (np. w ramach metody FIML)
- wiedza spoza próby statystycznej jest wprowadzana w sposób systematyczny, a nie nieformalny
- czasami mamy w modelu więcej zmiennych obserwowalnych niż wstrząsów – przy metodach klasycznych to problem

Dlaczego metody bayesowskie w DSGE?

Rozwiązany zlinearyzowany model DSGE jest przecież postaci $y_t = \mathrm{M}y_{t-1} + \mathrm{N}f_t$ i można go szacować metodą największej wiarygodności tak jak model SVAR...

- silne restrykcje teoretyczne dla ekonomicznie interpretowalnych wartości parametrów
- ucieczka parametrów w bezsensowne ekonomicznie zakresy grozi złamaniem warunków Blancharda-Kahna
- w badaniach empirycznych często dość krótkie próby – potrzeba informacji *a priori*
- model DSGE jest na ogół zbyt silnie wyidealizowany (zawiera zbyt mało sprzężeń w porównaniu z rzeczywistością), aby bez "wsparcia" *a priori* być skonfrontowanym z danymi (np. w ramach metody FIML)
- wiedza spoza próby statystycznej jest wprowadzana w sposób systematyczny, a nie nieformalny
- czasami mamy w modelu więcej zmiennych obserwowalnych niż wstrząsów – przy metodach klasycznych to problem

Dlaczego metody bayesowskie w DSGE?

Rozwiązany zlinearyzowany model DSGE jest przecież postaci $y_t = My_{t-1} + Nf_t$ i można go szacować metodą największej wiarygodności tak jak model SVAR...

- silne restrykcje teoretyczne dla ekonomicznie interpretowalnych wartości parametrów
- ucieczka parametrów w bezsensowne ekonomicznie zakresy grozi złamaniem warunków Blancharda-Kahna
- w badaniach empirycznych często dość krótkie próby – potrzeba informacji *a priori*
- model DSGE jest na ogół zbyt silnie wyidealizowany (zawiera zbyt mało sprzężeń w porównaniu z rzeczywistością), aby bez "wsparcia" *a priori* być skonfrontowanym z danymi (np. w ramach metody FIML)
- wiedza spoza próby statystycznej jest wprowadzana w sposób systematyczny, a nie nieformalny
- czasami mamy w modelu więcej zmiennych obserwowalnych niż wstrząsów – przy metodach klasycznych to problem

Dlaczego metody bayesowskie w DSGE?

Rozwiązany zlinearyzowany model DSGE jest przecież postaci $y_t = My_{t-1} + Nf_t$ i można go szacować metodą największej wiarygodności tak jak model SVAR...

- silne restrykcje teoretyczne dla ekonomicznie interpretowalnych wartości parametrów
- ucieczka parametrów w bezsensowne ekonomicznie zakresy grozi złamaniem warunków Blancharda-Kahna
- w badaniach empirycznych często dość krótkie próby – potrzeba informacji *a priori*
- model DSGE jest na ogół zbyt silnie wyidealizowany (zawiera zbyt mało sprzężeń w porównaniu z rzeczywistością), aby bez "wsparcia" *a priori* być skonfrontowanym z danymi (np. w ramach metody FIML)
- wiedza spoza próby statystycznej jest wprowadzana w sposób systematyczny, a nie nieformalny
- czasami mamy w modelu więcej zmiennych obserwowalnych niż wstrząsów – przy metodach klasycznych to problem

Podstawowy wzór ekonometrii bayesowskiej

Łączne prawdopodobieństwo danych i parametrów:

$$P(\mathrm{Y}, \boldsymbol{\theta}) = \begin{cases} P(\mathrm{Y}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta}) \\ P(\boldsymbol{\theta}|\mathrm{Y}) \cdot P(\mathrm{Y}) \end{cases}$$

Stąd reguła Bayesa:

$$P(\boldsymbol{\theta}|\mathrm{Y}) = \frac{P(\mathrm{Y}|\boldsymbol{\theta})}{P(\mathrm{Y})} \cdot P(\boldsymbol{\theta})$$

Innymi słowy, traktujemy tu wektor parametrów $\boldsymbol{\theta}$ jak wielowymiarową zmienną losową. Jej rozkład warunkowy względem danych $P(\boldsymbol{\theta}|\mathrm{Y})$ zależy od rozkładu bezwarunkowego parametrów $P(\boldsymbol{\theta})$ oraz rozkładu danych warunkowego względem parametrów $P(\mathrm{Y}|\boldsymbol{\theta})$.

Rozkład łączny *a posteriori*

$$P(\boldsymbol{\theta}|\mathrm{Y}) = \frac{P(\mathrm{Y}|\boldsymbol{\theta})}{P(\mathrm{Y})} \cdot P(\boldsymbol{\theta})$$

- $P(\boldsymbol{\theta})$ to gęstość *a priori*, obrazująca wiedzę o parametrze posiadaną zanim przystąpiliśmy do estymacji. O tym rozkładzie decydujemy sami w taki sposób, by odzwierciedlić nasze założenia, np. parametr ma należeć do przedziału od 0 do 1 (czyli przypisujemy prawdopodobieństwo *a priori* zero innym wartościom rzeczywistym) albo parametr ma wynosić około 0,6 (wówczas ustalamy taką średnią i niską wariancję wokół niej).
- $P(\mathrm{Y}|\theta)$ to gęstość próbkowa (identyczna z funkcją wiarygodności wykorzystywaną w metodach ML, FIML itp.).
- $P(\mathrm{Y})$ często pomijamy jako stałą skalującą (gwarantującą że funkcja gęstości całkuje się do 1), zapisując proporcjonalność $P(\boldsymbol{\theta}|\mathrm{Y}) \propto P(\mathrm{Y}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})$.

Rozkład łączny *a posteriori*

$$P(\boldsymbol{\theta}|\mathrm{Y}) = \frac{P(\mathrm{Y}|\boldsymbol{\theta})}{P(\mathrm{Y})} \cdot P(\boldsymbol{\theta})$$

- $P(\boldsymbol{\theta})$ to gęstość *a priori*, obrazująca wiedzę o parametrze posiadaną zanim przystąpiliśmy do estymacji. O tym rozkładzie decydujemy sami w taki sposób, by odzwierciedlić nasze założenia, np. parametr ma należeć do przedziału od 0 do 1 (czyli przypisujemy prawdopodobieństwo *a priori* zero innym wartościom rzeczywistym) albo parametr ma wynosić około 0,6 (wówczas ustalamy taką średnią i niską wariancję wokół niej).
- $P(\mathrm{Y}|\boldsymbol{\theta})$ to gęstość próbkowa (identyczna z funkcją wiarygodności wykorzystywaną w metodach ML, FIML itp.).
- $P(\mathrm{Y})$ często pomijamy jako stałą skalującą (gwarantującą że funkcja gęstości całkuje się do 1), zapisując proporcjonalność $P(\boldsymbol{\theta}|\mathrm{Y}) \propto P(\mathrm{Y}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})$.

Rozkład łączny *a posteriori*

$$P(\boldsymbol{\theta}|\mathrm{Y}) = \frac{P(\mathrm{Y}|\boldsymbol{\theta})}{P(\mathrm{Y})} \cdot P(\boldsymbol{\theta})$$

- $P(\boldsymbol{\theta})$ to gęstość *a priori*, obrazująca wiedzę o parametrze posiadaną zanim przystąpiliśmy do estymacji. O tym rozkładzie decydujemy sami w taki sposób, by odzwierciedlić nasze założenia, np. parametr ma należeć do przedziału od 0 do 1 (czyli przypisujemy prawdopodobieństwo *a priori* zero innym wartościom rzeczywistym) albo parametr ma wynosić około 0,6 (wówczas ustalamy taką średnią i niską wariancję wokół niej).
- $P(\mathrm{Y}|\boldsymbol{\theta})$ to gęstość próbkowa (identyczna z funkcją wiarygodności wykorzystywaną w metodach ML, FIML itp.).
- $P(\mathrm{Y})$ często pomijamy jako stałą skalującą (gwarantującą że funkcja gęstości całkuje się do 1), zapisując proporcjonalność $P(\boldsymbol{\theta}|\mathrm{Y}) \propto P(\mathrm{Y}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})$.

Rozkład brzegowy *a posteriori*

Łączny rozkład a posteriori ma w modelu DSGE wiele wymiarów (tyle ile parametrów), na ogół więcej niż 2, co uniemożliwia jego prezentację i interpretację.

W praktyce interesują nas raczej rozkłady brzegowe ze względu na pojedynczy parametr θ_i, tzn. rozkład łączny przecałkowany po wszystkich parametrach z wyjątkiem jednego. Np. przy trzech parametrach $\theta_1, \theta_2, \theta_3$:

$$P(\theta_1|\mathrm{Y}) \propto \iint P(\mathrm{Y}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})\, d\theta_2 d\theta_3$$

Nie możemy liczyć na istnienie analitycznych rozwiązań powyższego problemu – musimy numerycznie analizować rozkład łączny, przybliżając histogramy obrazujące rozkłady brzegowe *a posteriori*.

Metody bayesowskie – wyzwania

Estymacja bayesowska modelu DSGE wiąże się zatem z następującymi wyzwaniami dla ekonometryka:

1. wybór gęstości *a priori* $P(\theta)$ dla poszczególnych parametrów
2. konstrukcja funkcji wiarygodności $P(Y|\theta)$
3. określenie rozkładu *a posteriori*: łącznego i brzegowego dla pojedynczego parametru

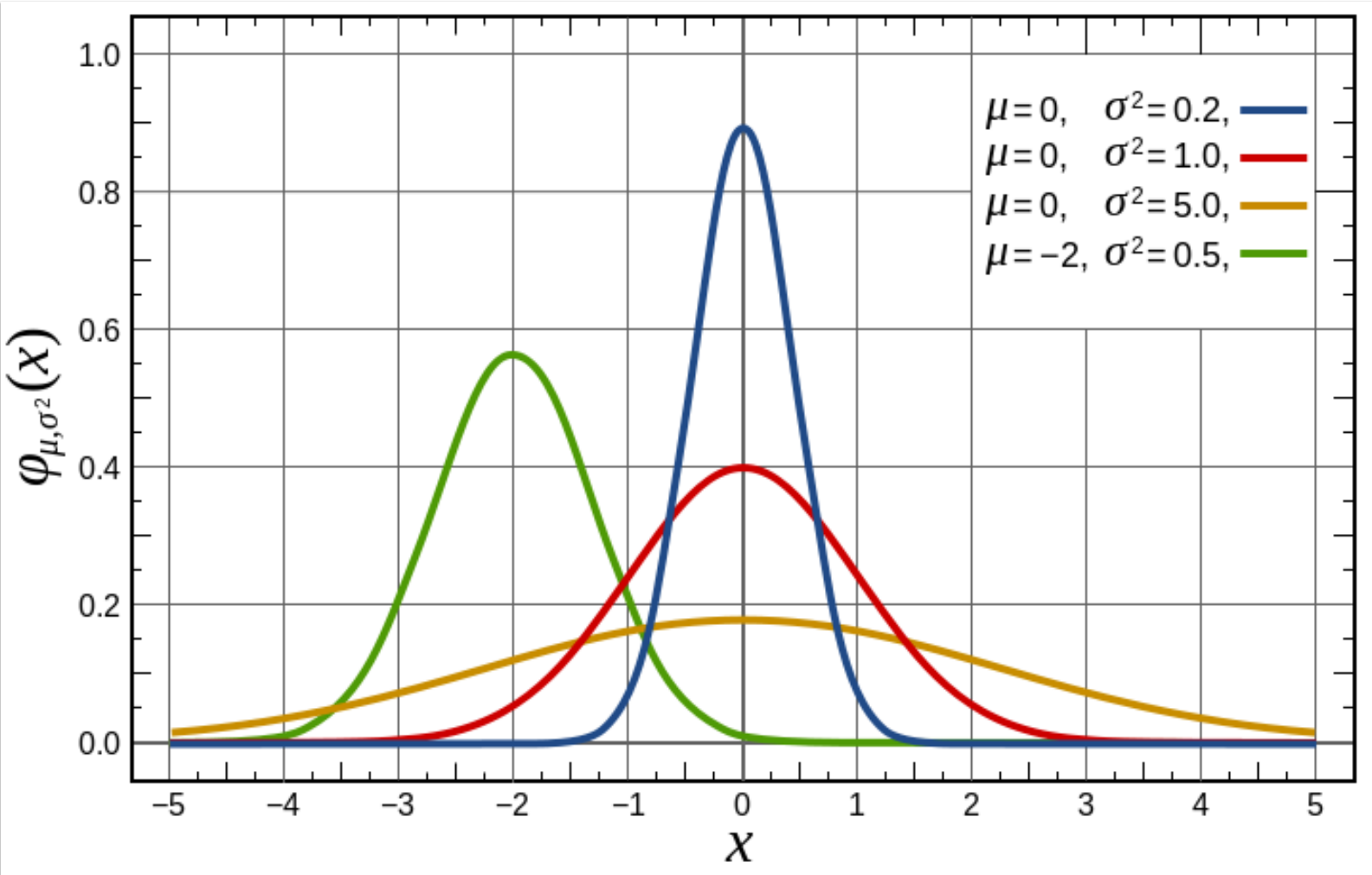
Zajmijmy się tymi tematami po kolei.

Plan prezentacji

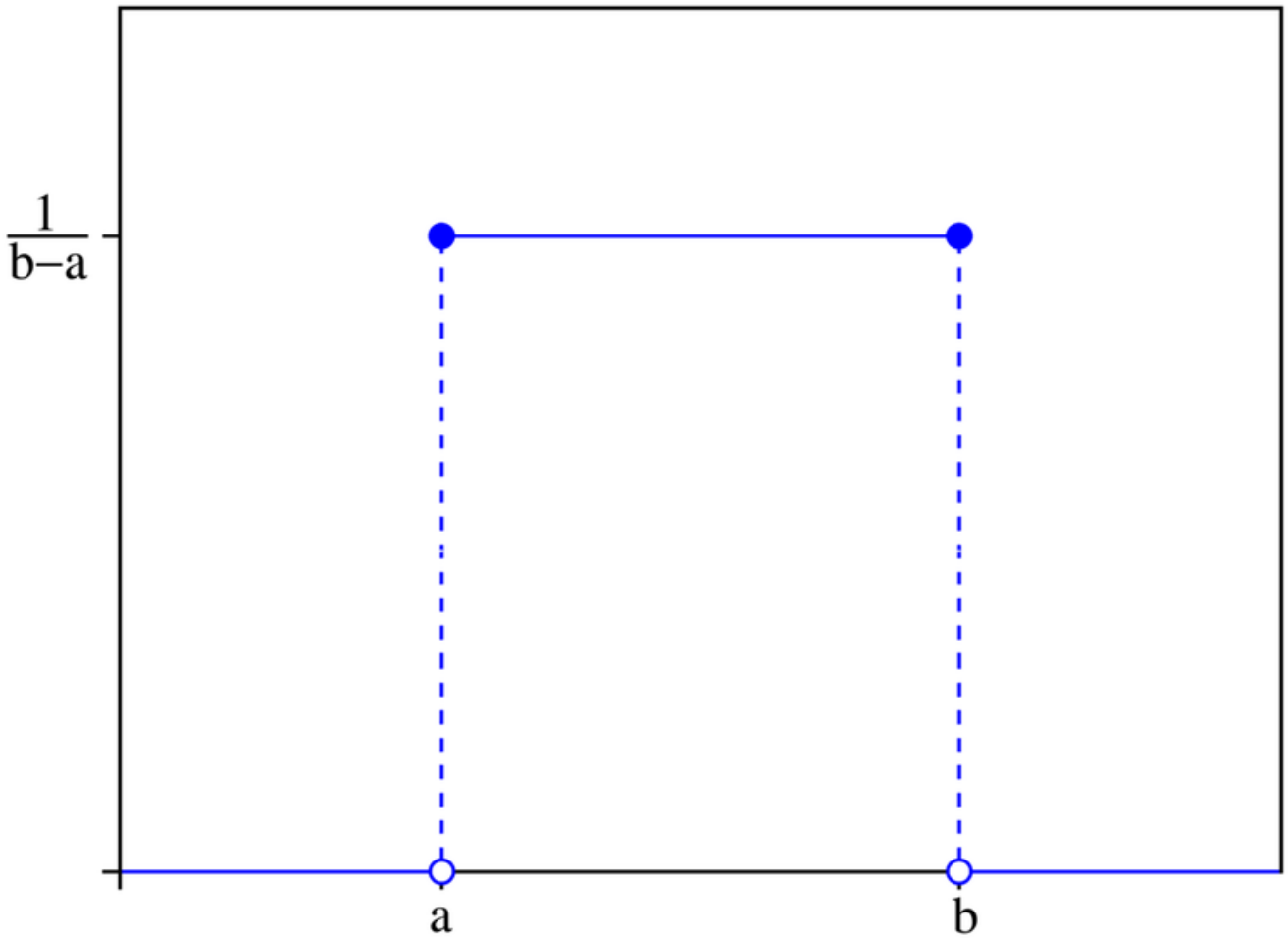
Typy rozkładów *a priori*

PRIOR_SHAPE	Corresponding distribution	Range
NORMAL_PDF	$N(\mu, \sigma)$	$\mathbb{R}$
GAMMA_PDF	$G_2(\mu, \sigma, p_3)$	$[p_3, +\infty)$
BETA_PDF	$B(\mu, \sigma, p_3, p_4)$	$[p_3, p_4]$
INV_GAMMA_PDF	$IG_1(\mu, \sigma)$	$\mathbb{R}^+$
UNIFORM_PDF	$U(p_3, p_4)$	$[p_3, p_4]$

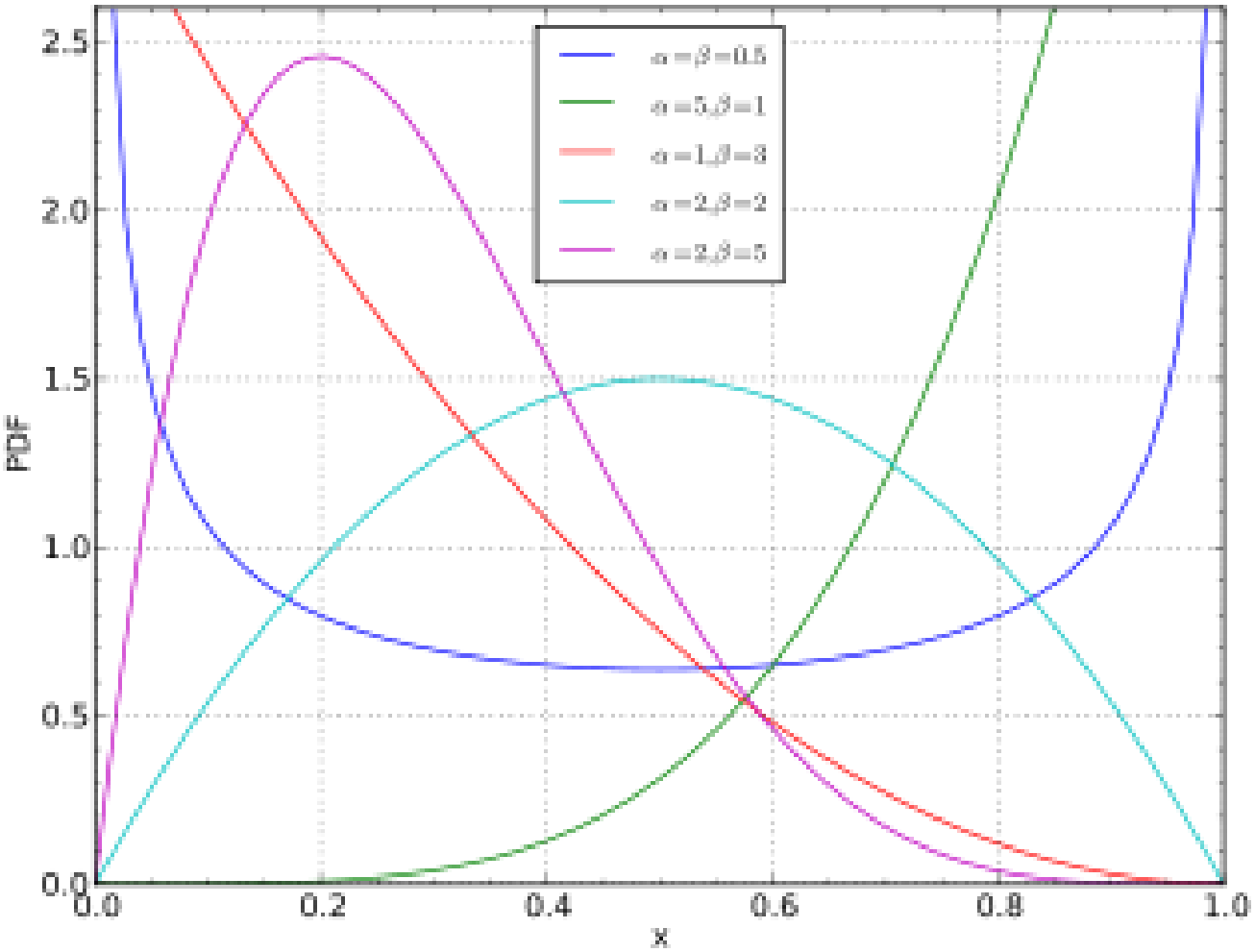
Rozkład normalny



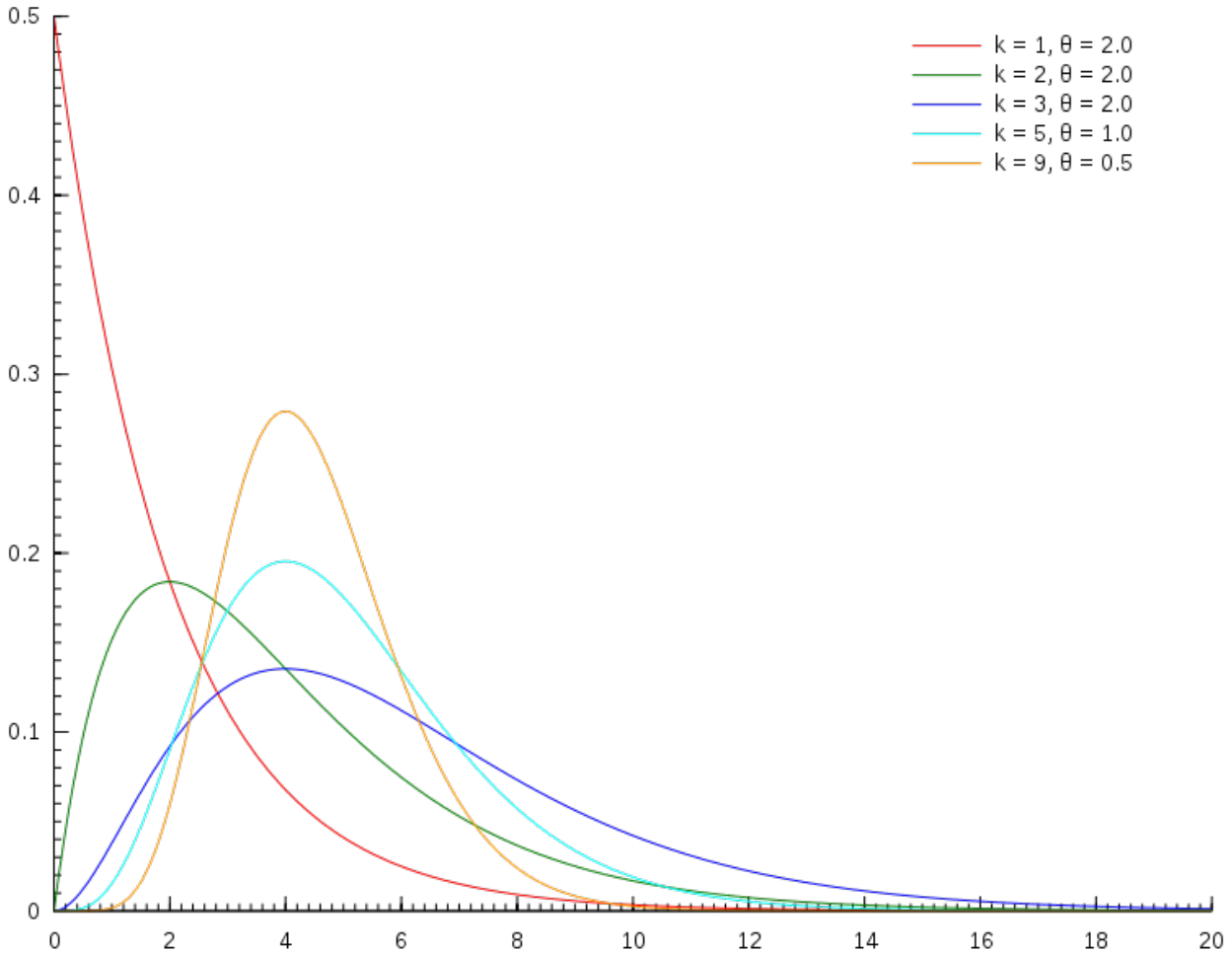
Rozkład jednostajny



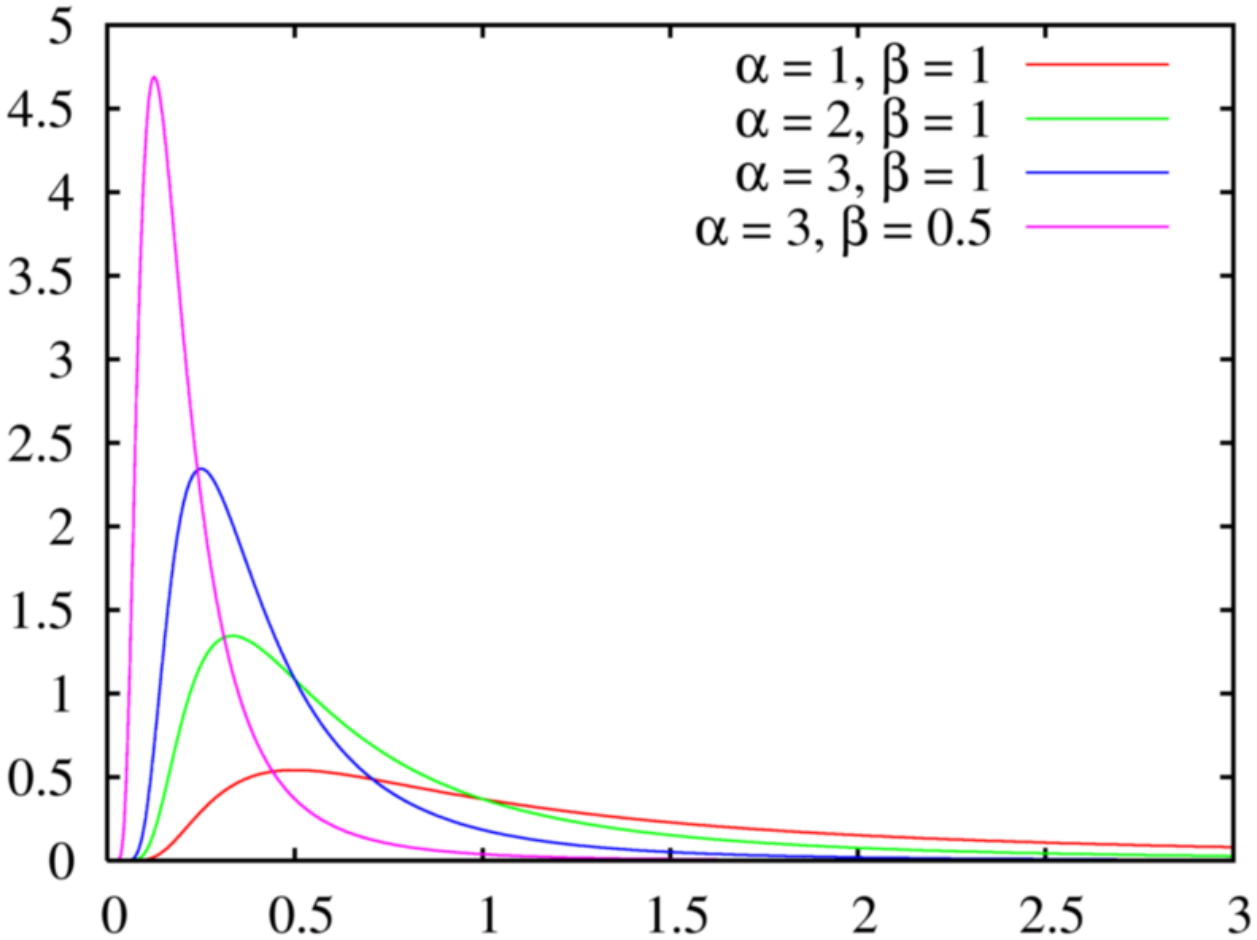
Rozkład beta



Rozkład gamma



Odwrotny rozkład gamma



Rozkłady *a priori* w Dynare

Blok **estimated_params** służy do określenia, które z parametrów są przedmiotem estymacji i określenia rozkładów *a priori*. Parametry, których tu nie ma, zostaną uznane za znane i skalibrowane w bloku parameters:

```
estimated_params;
h, beta_pdf, 0.7, 0.1;
sigma, gamma_pdf, 1.5, 0.4;
correl_d, uniform_pdf, , , -0.99,0.99;
sigma_d, inv_gamma_pdf, 4, inf;
end;
```

Plan prezentacji

Filtr Kalmana – idea

- Zmienne w linearyzowanym modelu DSGE wchodzą do modelu w postaci odchyleń od stanu ustalonego. W takiej sytuacji jest bardzo prawdopodobne, że nie mierzymy ich dokładnie, bo nie znamy stanu ustalonego.
  - Co więcej, nawet PKB czy konsumpcja nie są mierzone stuprocentowo dokładnie (a jedynie szacowane) i błędy pomiaru mogą rzutować na dopasowanie naszej misternej sieci zależności do danych.
  - W danych uwzględnione są także zajwiska, których brak w naszym modelu. One nie zawsze pasują do naszych "strukturalnych" wstrząsów, które wbudowujemy w reguły deycyzjne podmiotów.

- Pojawia się zatem problem „ekstrakcji sygnału”, czyli identyfikacji i pomiaru zmiennej nieobserwowalnej na podstawie jej charakterystyki ekonomicznej, jak również sieci zależności między nią a obserwowalnymi.
  - Tę sieć zależności opisuje tzw. **model w przestrzeni stanów**.
- Przy estymacji pojawia się jednak trudność ze skonstruowaniem funkcji wiarygodności, bo występują w niej zmienne nieobserwowalne.
  - Aby przezwyciężyć problem, musimy posłużyć się tzw. **filtrem Kalmana**.

Model w przestrzeni stanów

Liniowy, Gaussowski:

$$\begin{cases} \boldsymbol{\alpha_t} &= \boldsymbol{\mu} + \mathrm{F}\boldsymbol{\alpha_{t-1}} + \boldsymbol{\varepsilon_t} \\ \boldsymbol{y_t} &= \mathrm{H}\boldsymbol{\alpha_t} + \mathrm{Ax_t} + \boldsymbol{u_t} \end{cases}$$

ZMIENNA STANU:

- $\boldsymbol{\alpha_t}$ – wektor $m \times 1$ zmiennych modelu (nieobserwowalnych)
- $\boldsymbol{\varepsilon_t} \sim N(0, \mathrm{Q})$– wstrząsy w modelu

ZMIENNA SYGNAŁU:

- $\boldsymbol{y_t}$ – wektor $n \times 1$ zmiennych obserwowalnych
- $\boldsymbol{u}_t \sim N(0, \mathrm{R})$ – błędy pomiaru

Rozkłady zmiennych stanu

$$\boldsymbol{\alpha_t} = \boldsymbol{\mu} + \mathsf{F}\boldsymbol{\alpha_{t-1}} + \boldsymbol{\varepsilon_t}$$

- Przyjmujemy, że wektor $\boldsymbol{\alpha_{t-1}}$ jest zmienną losową o wielowymiarowym rozkładzie normalnym (MVN) z wartością oczekiwaną $\boldsymbol{a_{t-1}}$ oraz wariancją $\boldsymbol{P_{t-1}}$.
- Wówczas $\boldsymbol{\alpha_t}$ – jako kombinacja dwóch zmiennych o rozkładzie MVN – też ma rozkładm MVN.
- Pozostaje pytanie o rozkład $\boldsymbol{\alpha_1}$ (rekurencyjna definicja sprawia, że nie będziemy znali $\boldsymbol{a_0}$ ani $\boldsymbol{P_0}$).

Problem inicjalizacji

Przed okresem t znamy:

- y_{t-1}
- a_{t-1} – wartość oczekiwana α_{t-1} warunkowa względem y_{t-1}
- P_{t-1} – macierz wariancji-kowariancji α_{t-1} warunkowa względem y_{t-1}

Problem inicjalizacji: a_0, P_0 – zwykle długookresowe wartości bezwarunkowe (***Hamilton, 1994***):

- $a_0 = \mu$
- $vec P_0 = [I_{m^2} - (F \otimes F)]^{-1} \cdot vec(Q)$

Możliwe również przyjęcie alternatywnego założenia: że α_0 ma tzw. rozkład nieinformacyjny (b. wysoka wariancja). Wówczas wyprowadzenie staje się nieco bardziej skomplikowane niż prezentowane tutaj.

Równania predykcji stanów

Co wiemy *ex ante* (przed t) na temat stanów w t?

$$E_{t-1}(\boldsymbol{\alpha}_t) = \boldsymbol{F} \cdot E_{t-1}(\boldsymbol{\alpha}_{t-1}) + \boldsymbol{\mu} = \mathrm{F}\mathrm{a}_{t-1} + \boldsymbol{\mu} \equiv \boldsymbol{a}_{t|t-1}$$

$$\begin{aligned}
&E_{t-1}\left[\left(\boldsymbol{\alpha}_t - \boldsymbol{a}_{t|t-1}\right)\left(\boldsymbol{\alpha}_t - \boldsymbol{a}_{t|t-1}\right)^T\right] = \\
&= E_{t-1}\left[\left(\boldsymbol{F}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t - \boldsymbol{F}\mathrm{a}_{t-1} - \boldsymbol{\mu}\right)\left(\boldsymbol{F}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t - \boldsymbol{F}\mathrm{a}_{t-1} - \boldsymbol{\mu}\right)^T\right] = \\
&= E_{t-1}\left\{\left[\boldsymbol{F}\left(\boldsymbol{\alpha}_{t-1} - \mathrm{a}_{t-1}\right) + \boldsymbol{\varepsilon}_t\right]\left[\boldsymbol{F}\left(\boldsymbol{\alpha}_{t-1} - \mathrm{a}_{t-1}\right) + \boldsymbol{\varepsilon}_t\right]^T\right\} = \\
&= \mathrm{F}\mathrm{P}_{t-1}\mathrm{F}^T + \mathrm{Q} \equiv \mathrm{P}_{t|t-1}
\end{aligned}$$

Równania predykcji obserwacji

Co wiemy *ex ante* (przed t) na temat obserwacji w t?

- $E_{t-1}(\boldsymbol{y}_t) = \mathrm{H}\boldsymbol{a}_{t|t-1} + \mathrm{Ax_t} \equiv \boldsymbol{y}_{t|t-1}$
- $E_{t-1}\left\{[\boldsymbol{y}_t - E_{t-1}(\boldsymbol{y}_t)][\boldsymbol{y}_t - E_{t-1}(\boldsymbol{y}_t)]^T\right\} = \mathrm{H}\boldsymbol{P}_{t|t-1}\mathrm{H}^T + \mathrm{R} \equiv \boldsymbol{V}_t$

Kowariancja *ex ante* stanu i obserwacji

$$
\begin{aligned}
&E_{t-1}\left\{\left[\alpha_t - \boldsymbol{a}_{t|t-1}\right]\left[\boldsymbol{y}_t - \boldsymbol{H}\boldsymbol{a}_{t|t-1} - \boldsymbol{A}\boldsymbol{x}_t\right]^T\right\} = \\
&= E_{t-1}\left\{\left[\alpha_t - \boldsymbol{a}_{t|t-1}\right]\left[\boldsymbol{H}\alpha_t + \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{u}_t - \boldsymbol{H}\boldsymbol{a}_{t|t-1} - \boldsymbol{A}\boldsymbol{x}_t\right]^T\right\} = \\
&= E_{t-1}\left\{\left[\alpha_t - \boldsymbol{a}_{t|t-1}\right]\left[\alpha_t - \boldsymbol{a}_{t|t-1}\right]^T \boldsymbol{H}^T\right\} = \boldsymbol{P}_{t|t-1}\boldsymbol{H}^T
\end{aligned}
$$

Ta kowariancja uzupełnia opis łącznego rozkładu *ex ante* α_t i $\boldsymbol{y}_t$ jako $MVN\left(\begin{bmatrix} \boldsymbol{a}_{t|t-1} \\ \boldsymbol{y}_{t|t-1} \end{bmatrix}, \begin{bmatrix} \boldsymbol{P}_{t|t-1} & \boldsymbol{P}_{t|t-1}\boldsymbol{H}^T \\ \boldsymbol{H}\boldsymbol{P}^T_{t|t-1} & \boldsymbol{V}_t \end{bmatrix}\right)$.

Równania aktualizacji

Czego się dowiedzieliśmy po zaobserwowaniu $\boldsymbol{y}_t$ (*ex post* w t):

- $\boldsymbol{e}_t = \boldsymbol{y}_t - \boldsymbol{y}_{t|t-1}$

Z perspektywy okresu t, $\boldsymbol{y}_t$ nie jest już zmienną losową, a więc rozkład $\boldsymbol{\alpha}_t$ wyznaczamy na podstawie własności MVN jako rozkład warunkowy:

- $E_t(\boldsymbol{\alpha}_t) = \boldsymbol{a}_{t|t-1} + \boldsymbol{P}_{t|t-1}\mathrm{H}^T\mathrm{V}_t^{-1}\boldsymbol{e}_t \equiv \boldsymbol{a}_t$
- $E_t\left[(\boldsymbol{\alpha}_t - \boldsymbol{a}_t)(\boldsymbol{\alpha}_t - \boldsymbol{a}_t)^T\right] = \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\mathrm{H}^T\boldsymbol{V}_t^{-1}\mathrm{H}\boldsymbol{P}_{t|t-1} \equiv \boldsymbol{P}_t$

Funkcja wiarygodności

$$\mathcal{L}\left[\mathrm{y}, \boldsymbol{\theta}\right] = -\frac{Tn}{2} ln\left(2\pi\right) + \frac{T}{2}\sum_{t=1}^{T} ln\left|\mathrm{V}_t\left(\boldsymbol{\theta}\right)\right| - \frac{1}{2}\sum_{t=1}^{T} \boldsymbol{e}_t\left(\boldsymbol{\theta}\right)^T \mathrm{V}_t\left(\boldsymbol{\theta}\right)^{-1} \boldsymbol{e}_t\left(\boldsymbol{\theta}\right)$$

$$\rightarrow \max_{\boldsymbol{\theta}}$$

Formy prezentacji zmiennych nieobserwowalnych

- Wartości zmiennych nieobserwowalnych (stanów) mogą być przybliżane punktowo poprzez $\hat{\boldsymbol{a}}_t$ (po oszacowaniu macierzy parametrów modelu)...
- ...lub przedziałowo, z dodatkowym wykorzystaniem $\hat{\boldsymbol{P}}_t$ (rozkład MVN dla $\boldsymbol{\alpha}_t$).
- Należy jednak pamiętać, że oba te wektory (tzw. zmienne **filtrowane**) bazują na wiedzy od okresu 1 do t.
- Ekonometryk dysponujący wiedzą do T włącznie jest w stanie dodatkowo wyznaczyć wartości **wygładzone** zmiennych nieobserwowalnych (wzory: zob. ***Hamilton, 1994***, roz. 13.6).

Model przestrzeni stanów i filtr Kalmana w Dynare

Dynare automatycznie określa model przestrzeni stanów i konstruuje funkcję wiarygodności, zakładając przy tym, że:

1. wszystkie nasze zmienne w pliku .mod to zmienne nieobserwowalne (stany, α_t)
2. mają one swoje "obserwowalne" odpowiedniki, a więc są mierzone w sposób niedoskonały
   1. macierz Z jest jednostkowa
   2. równania stanu to równania naszego modelu
3. wstrząsy strukturalne oraz błędy pomiaru mają rozkład normalny

Zmienne obserwowalne w Dynare

Po bloku model w pliku .mod wpisujemy blok **varobs**, czyli które ze zmiennych endogenicznych są obserwowalne i będą wykorzystane w naszej analizie empirycznej:

```
varobs y i pi;
```

Dane muszą zostać dostarczone w pliku przestrzeni roboczej .mat (zob. wcześniejsze zajęcia). Plik ten musi zawierać pionowe wektory z danymi o jednakowej długości, nazywające się tak samo, jak nasze zmienne w bloku **varobs**.

Błędy pomiaru w Dynare

Jeżeli wprowadzamy błędy pomiaru w filtrze Kalmana, to dodatkowo w bloku **estimated_params** wpisujemy wiersze postaci:

```
stderr y, inv_gamma_pdf, 0.1, inf;
stderr ir, inv_gamma_pdf, 0.1, inf;
stderr pi, inv_gamma_pdf, 0.1, inf;
```

Plan prezentacji

Algorytm Metropolisa-Hastingsa – przypadek ogólny

Rozważmy wektor parametrów $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_K)$ o nieznanej gęstości *a posteriori* $p(\boldsymbol{\theta}|\boldsymbol{y})$. Nie możemy również ustalić **rozkładów warunkowych**.

1. Wybieramy wektor startowy $\boldsymbol{\theta}^{(0)} = (\theta_1, \theta_2, ..., \theta_K)$.
2. Losujemy „kandydata" θ^* korzystając z gęstości generującej kandydatów, $q\left(\theta|\theta^{(0)}\right)$.
3. Obliczamy prawdopodobieństwo akceptacji kandydata, $\alpha\left(\theta^*, \theta^{(0)}\right)$.
4. $$\theta^{(1)} = \begin{cases} \theta^* & z\ prawdopodobieństwem\ \alpha\left(\theta^*, \theta^{(0)}\right) \\ \theta^{(0)} & z\ prawdopodobieństwem\ 1 - \alpha\left(\theta^*, \theta^{(0)}\right) \end{cases}$$
5. Powtarzamy kroki 1-4 z wektorem startowym $\theta^{(1)}$.
6. Powtarzamy sekwencję S razy.

Algorytm Metropolisa-Hastingsa – przypadek ogólny

Rozważmy wektor parametrów $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_K)$ o nieznanej gęstości *a posteriori* $p(\boldsymbol{\theta}|\boldsymbol{y})$. Nie możemy również ustalić **rozkładów warunkowych**.

1. Wybieramy wektor startowy $\boldsymbol{\theta}^{(0)} = (\theta_1, \theta_2, ..., \theta_K)$.
2. Losujemy „kandydata" $\boldsymbol{\theta}^*$ korzystając z gęstości generującej kandydatów, $q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(0)}\right)$.
3. Obliczamy prawdopodobieństwo akceptacji kandydata, $\alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right)$.
4. $\boldsymbol{\theta}^{(1)} = \begin{cases} \boldsymbol{\theta}^* & z\ prawdopodobieństwem\ \alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \\ \boldsymbol{\theta}^{(0)} & z\ prawdopodobieństwem\ 1 - \alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \end{cases}$
5. Powtarzamy kroki 1-4 z wektorem startowym $\boldsymbol{\theta}^{(1)}$.
6. Powtarzamy sekwencję S razy.

Algorytm Metropolisa-Hastingsa – przypadek ogólny

Rozważmy wektor parametrów $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_K)$ o nieznanej gęstości *a posteriori* $p(\boldsymbol{\theta}|\boldsymbol{y})$. Nie możemy również ustalić **rozkładów warunkowych**.

1. Wybieramy wektor startowy $\boldsymbol{\theta}^{(0)} = (\theta_1, \theta_2, ..., \theta_K)$.
2. Losujemy „kandydata" $\boldsymbol{\theta}^*$ korzystając z gęstości generującej kandydatów, $q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(0)}\right)$.
3. Obliczamy prawdopodobieństwo akceptacji kandydata, $\alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right)$.
4. $\theta^{(1)} = \begin{cases} \theta^* & z\ prawdopodobieństwem\ \alpha\left(\theta^*, \theta^{(0)}\right) \\ \theta^{(0)} & z\ prawdopodobieństwem\ 1-\alpha\left(\theta^*, \theta^{(0)}\right) \end{cases}$
5. Powtarzamy kroki 1-4 z wektorem startowym $\theta^{(1)}$.
6. Powtarzamy sekwencję S razy.

Algorytm Metropolisa-Hastingsa – przypadek ogólny

Rozważmy wektor parametrów $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_K)$ o nieznanej gęstości *a posteriori* $p(\boldsymbol{\theta}|\boldsymbol{y})$. Nie możemy również ustalić **rozkładów warunkowych**.

1. Wybieramy wektor startowy $\boldsymbol{\theta}^{(0)} = (\theta_1, \theta_2, ..., \theta_K)$.
2. Losujemy „kandydata" $\boldsymbol{\theta}^*$ korzystając z gęstości generującej kandydatów, $q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(0)}\right)$.
3. Obliczamy prawdopodobieństwo akceptacji kandydata, $\alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right)$.
4. $$\boldsymbol{\theta}^{(1)} = \begin{cases} \boldsymbol{\theta}^* & z\ prawdopodobieństwem\ \alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \\ \boldsymbol{\theta}^{(0)} & z\ prawdopodobieństwem\ 1 - \alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \end{cases}$$
5. Powtarzamy kroki 1-4 z wektorem startowym $\boldsymbol{\theta}^{(1)}$.
6. Powtarzamy sekwencję S razy.

Algorytm Metropolisa-Hastingsa – przypadek ogólny

Rozważmy wektor parametrów $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_K)$ o nieznanej gęstości *a posteriori* $p(\boldsymbol{\theta}|\boldsymbol{y})$. Nie możemy również ustalić **rozkładów warunkowych**.

1. Wybieramy wektor startowy $\boldsymbol{\theta}^{(0)} = (\theta_1, \theta_2, ..., \theta_K)$.
2. Losujemy „kandydata" $\boldsymbol{\theta}^*$ korzystając z gęstości generującej kandydatów, $q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(0)}\right)$.
3. Obliczamy prawdopodobieństwo akceptacji kandydata, $\alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right)$.
4. $$\boldsymbol{\theta}^{(1)} = \begin{cases} \boldsymbol{\theta}^* & z\ prawdopodobieństwem\ \alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \\ \boldsymbol{\theta}^{(0)} & z\ prawdopodobieństwem\ 1 - \alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \end{cases}$$
5. Powtarzamy kroki 1-4 z wektorem startowym $\boldsymbol{\theta}^{(1)}$.
6. Powtarzamy sekwencję S razy.

Algorytm Metropolisa-Hastingsa – przypadek ogólny

Rozważmy wektor parametrów $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_K)$ o nieznanej gęstości *a posteriori* $p(\boldsymbol{\theta}|\boldsymbol{y})$. Nie możemy również ustalić **rozkładów warunkowych**.

1. Wybieramy wektor startowy $\boldsymbol{\theta}^{(0)} = (\theta_1, \theta_2, ..., \theta_K)$.
2. Losujemy „kandydata" $\boldsymbol{\theta}^*$ korzystając z gęstości generującej kandydatów, $q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(0)}\right)$.
3. Obliczamy prawdopodobieństwo akceptacji kandydata, $\alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right)$.
4. $$\boldsymbol{\theta}^{(1)} = \begin{cases} \boldsymbol{\theta}^* & z\ prawdopodobieństwem\ \alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \\ \boldsymbol{\theta}^{(0)} & z\ prawdopodobieństwem\ 1-\alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) \end{cases}$$
5. Powtarzamy kroki 1-4 z wektorem startowym $\boldsymbol{\theta}^{(1)}$.
6. Powtarzamy sekwencję S razy.

Algorytm MH – wybór wektora startowego

1. Arbitralne wskazanie przez użytkownika.
2. Losowanie z brzegowych rozkładów *a priori*, jeżeli możliwe.
3. Wielokrotne wybory z analizą wrażliwości.
   - W każdym przypadku odcinamy pewną liczbę początkowych wartości jako *burn-in*. Przy prawidłowej zbieżności oraz odpowiednio wysokiej liczbie iteracji (S) nie powinno to mieć znaczenia.

Algorytm MH – gęstość generująca kandydatów

1. W ogólnym przypadku zakładamy, że jest zależna od bieżącego punktu $\boldsymbol{\theta}^{(s)}$.
   - Najczęstszą implementacją jest ***Random Walk*** **MH**, gdzie losujemy $\varepsilon \sim N\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\Sigma}\right)$ i rozważamy kandydata:

$$\boldsymbol{\theta}^{*} = \boldsymbol{\theta}^{(s)} + \varepsilon$$

2. Nie musi tak jednak być. Wówczas mówimy o implementacji *Independence Chain* MH (nie omawiamy tutaj).

Algorytm MH – gęstość generująca kandydatów

1. W ogólnym przypadku zakładamy, że jest zależna od bieżącego punktu $\boldsymbol{\theta}^{(s)}$.
   - Najczęstszą implementacją jest ***Random Walk*** **MH**, gdzie losujemy $\boldsymbol{\varepsilon} \sim N\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\Sigma}\right)$ i rozważamy kandydata:
   $$\boldsymbol{\theta}^{*} = \boldsymbol{\theta}^{(s)} + \boldsymbol{\varepsilon}$$
2. Nie musi tak jednak być. Wówczas mówimy o implementacji ***Independence Chain*** **MH** (nie omawiamy tutaj).

Algorytm MH – prawdopodobieństwo akceptacji α

1. Ponieważ q to funkcja umowna, korzystanie z niej bez dodatkowych korekt nie gwarantuje nam uzyskania sekwencji losowań przybliżających rozkład *a posteriori*.
   1. Algorytm bez korekt zbyt często pozostaje w obszarach o wysokiej gęstości *a posteriori*.
   2. W związku z tym musimy go skorygować, by dostatecznie dobrze „zwiedzić" całą dziedzinę parametrów.
2. Korekta polega na nieakceptowaniu wszystkich kandydatów wylosowanych na podstawie gęstości q. W przypadku braku akceptacji, kolejnym elementem łańcucha jest kopia poprzedniego.
   - W implementacji ***Random Walk*** wzór na prawdopodobieństwo akceptacji zależy od gęstości *a posteriori* (p) dla wektorów: poprzedniego ($\theta^{(s)}$) oraz kandydata (θ^*). Szczegóły: Koop (roz. 5.5).

$$\alpha\left(\theta^*, \theta^{(0)}\right) = \min\left[\frac{p(\theta^*|y)}{p(\theta^{(s-1)}|y)}, 1\right]$$

Algorytm MH – prawdopodobieństwo akceptacji α

1. Ponieważ q to funkcja umowna, korzystanie z niej bez dodatkowych korekt nie gwarantuje nam uzyskania sekwencji losowań przybliżających rozkład *a posteriori*.
   1. Algorytm bez korekt zbyt często pozostaje w obszarach o wysokiej gęstości *a posteriori*.
   2. W związku z tym musimy go skorygować, by dostatecznie dobrze „zwiedzić" całą dziedzinę parametrów.
2. Korekta polega na nieakceptowaniu wszystkich kandydatów wylosowanych na podstawie gęstości q. W przypadku braku akceptacji, kolejnym elementem łańcucha jest kopia poprzedniego.
   - W implementacji ***Random Walk*** wzór na prawdopodobieństwo akceptacji zależy od gęstości *a posteriori* (p) dla wektorów: poprzedniego ($\boldsymbol{\theta}^{(s)}$) oraz kandydata ($\boldsymbol{\theta}^*$). Szczegóły: Koop (roz. 5.5).

$$\alpha\left(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(0)}\right) = \min\left[\frac{p\left(\boldsymbol{\theta}^*|\boldsymbol{y}\right)}{p\left(\boldsymbol{\theta}^{(s-1)}|\boldsymbol{y}\right)}, 1\right]$$

Algorytm MH – α versus q

- Miarą jakości wyników jest m.in. średnie prawdopodobieństwo akceptacji $\bar{\alpha}$.
- Okazuje się, że optymalne wartości $\bar{\alpha} \in [0,2;\ 0,4]$. Dostatecznie niskie prawdopodobieństwo akceptacji oznacza, że dziedzina gęstości *a posteriori* została dobrze wyeksplorowana.
- $\bar{\alpha}$ to jednak wartość wynikowa i nie możemy jej wprost wybrać. Zależy ona przede wszystkim od doboru gęstości generującej kandydatów q.
  - W przypadku ***Random Walk MH***, sprowadza się to do odpowiedniego ustalenia wariancji kroku ε, czyli $\boldsymbol{\Sigma}$(w Dynare: **mh_jscale**)
  - Relację między $\bar{\alpha}$ a $\boldsymbol{\Sigma}$ należy zbadać w ramach dodatkowej procedury iteracyjnej.
  - Zaczynamy w niej od $\boldsymbol{\Sigma}^{(0)} = c^{(0)} \cdot \boldsymbol{I}$. W przypadku zbyt wysokiego $\bar{\alpha}^{(0)}$zbyt często akceptujemy, a więc jesteśmy zbyt konserwatywni w „zwiedzaniu" dziedziny, czyli powinnismy ustalić $c^{(1)} > c^{(0)}$.

Algorytm MH w Dynare (1)

Polecenie estymacji jest skonstruowane w następujący sposób:

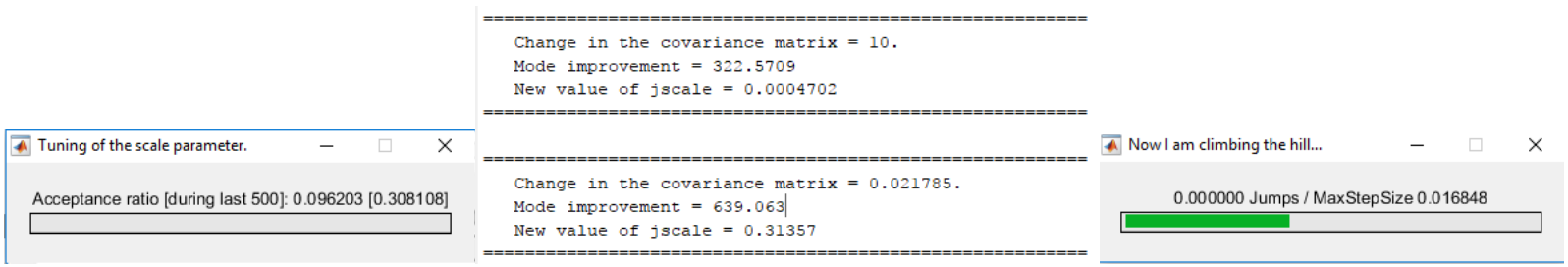
```
estimation(datafile = plik_danych, mh_replic = 20000,
mh_nblocks = 5, mh_drop = 0.5, mh_jscale = 0.2);
```

Kluczowe opcje:

- **plik_danych.mod** to plik w przestrzeni roboczej zawierający wektory zmiennych obserwowalnych;
- **mh_replic** – liczba replikacji algorytmu (długość łańcucha);
- **mh_nblocks** – liczba łańcuchów (co najmniej 2 – ale lepiej więcej, np. 5-10, tyle że to jest czasochłonne...);
- **mh_drop** – frakcja odrzuconych fragmentów łańcucha blisko początku (by uniezależnić się od losowo wybranego początku).
- dalsze opcje w manualu Dynare (np. dobór podpróby)

Algorytm MH w Dynare (2)

mh_jscale to parametr powiązany z wyborem odpowiedniej wariancji kroku i podlega ew. korekcie w przebiegu estymacji (następnym razem możemy podać inną wartość, aby przyspieszyć):



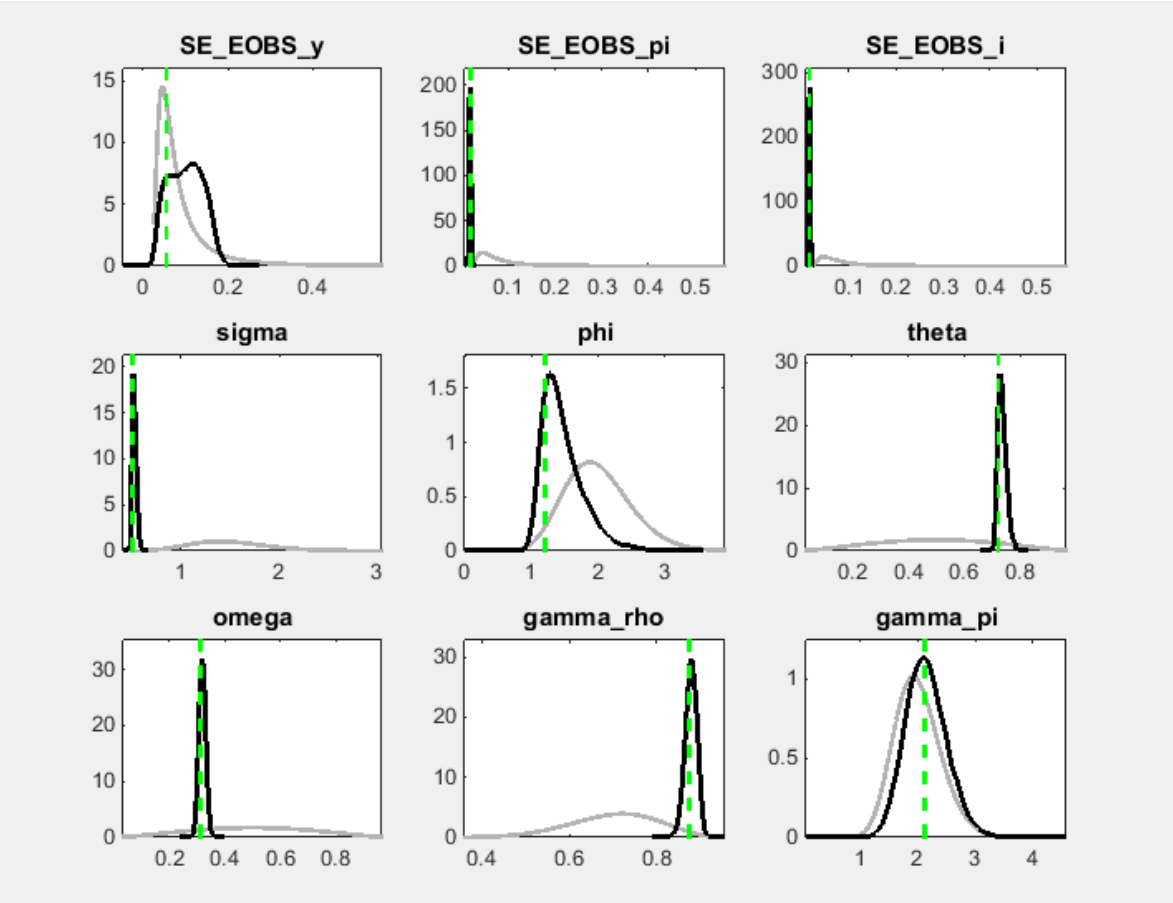
Jak ocenić wyniki estymacji? (1a)

- Rolę oszacowania punktowego pełni na ogół średnia z rozkładu *a posteriori*. Możemy zastanowić się nad jej wymiarem ekonomicznym i ulokowaniem na tle naszej wiedzy *a priori*.
- Statystyczny "sukces" osiągamy, gdy rozkład *a posteriori* widocznie różni się od rozkładu *a priori* bardziej smukłym kształtem (niższą wariancją) – oznacza to, że próba sprecyzowała naszą wiedzę o parametrze w stosunku do tej, którą posiadaliśmy wcześniej. Rozkłady *a posteriori* powinny mieć ponadto stosunkowo regularny kształt (np. być jednomodalne).

Jak ocenić wyniki estymacji? (1a)

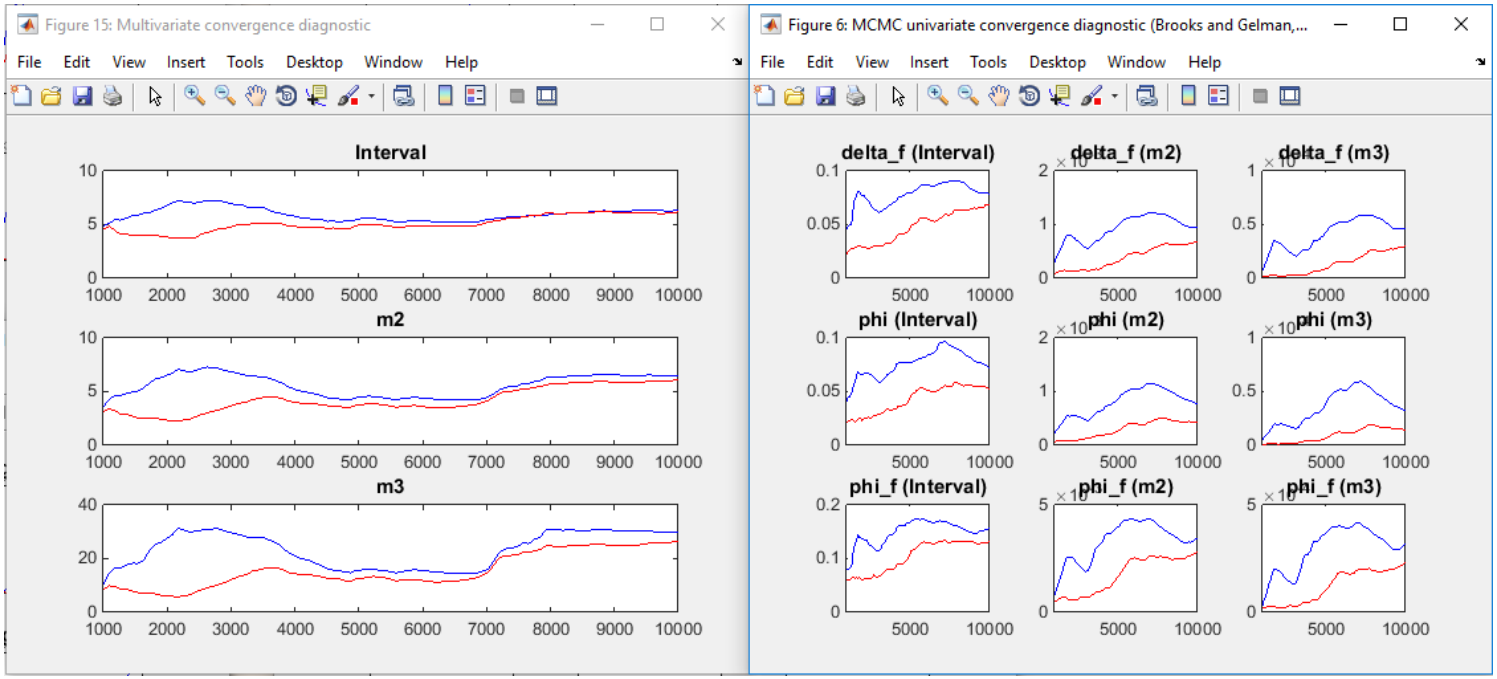
- Rolę oszacowania punktowego pełni na ogół średnia z rozkładu *a posteriori*. Możemy zastanowić się nad jej wymiarem ekonomicznym i ulokowaniem na tle naszej wiedzy *a priori*.
- Statystyczny "sukces" osiągamy, gdy rozkład *a posteriori* widocznie różni się od rozkładu *a priori* bardziej smukłym kształtem (niższą wariancją) – oznacza to, że próba sprecyzowała naszą wiedzę o parametrze w stosunku do tej, którą posiadaliśmy wcześniej. Rozkłady *a posteriori* powinny mieć ponadto stosunkowo regularny kształt (np. być jednomodalne).

Jak ocenić wyniki estymacji? (1b)



Jak ocenić wyniki estymacji? (2a)

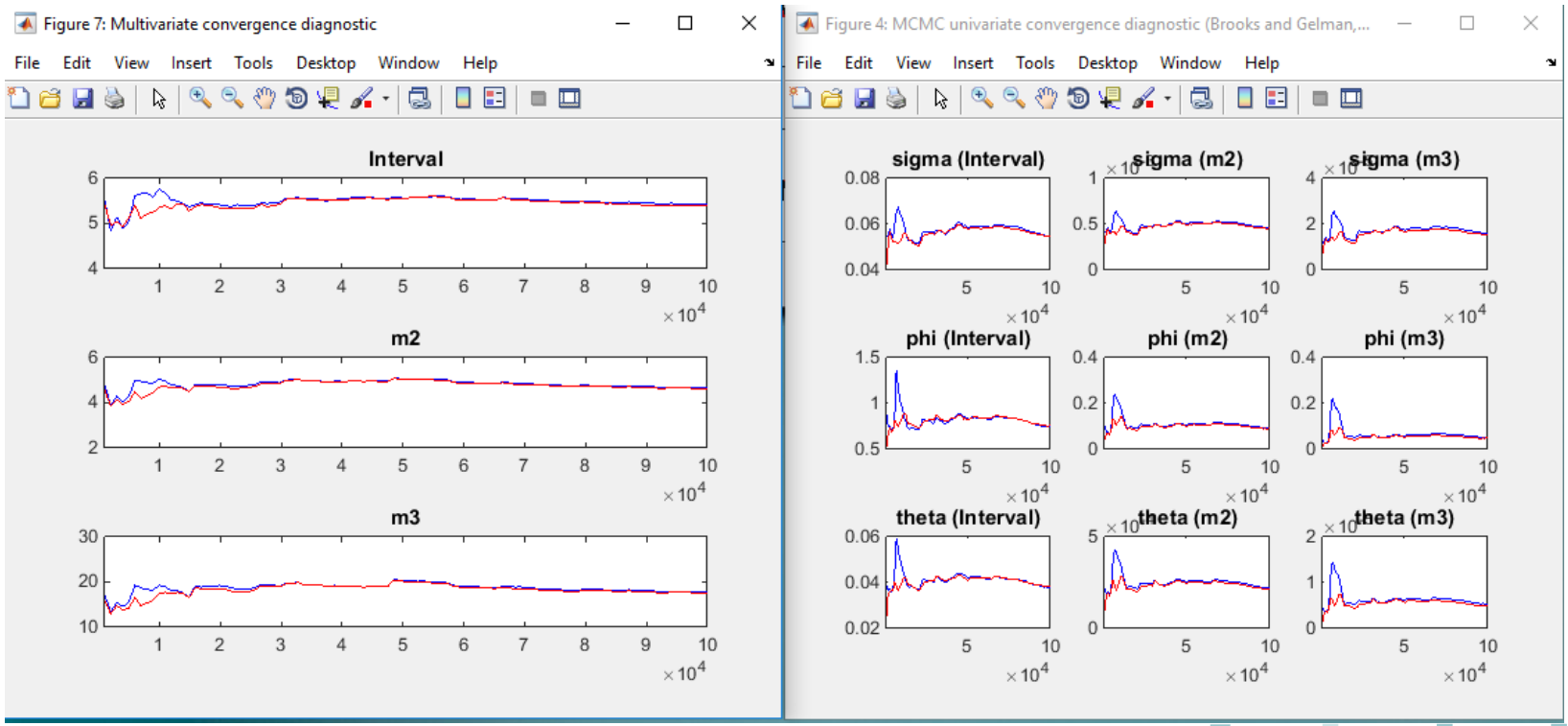
Czerwona linia: wariancja wewnątrzłańcuchowa (**within**). Niebieska linia: całkowita wariancja (**within + between**).



Jak ocenić wyniki estymacji? (2b)

Stabilizacja czerwonej linii: losowanie z tego samego rozkładu (symptom osiągnięcia zbieżności).

Zbieżność obu linii: brak wariancji międzyłańcuchowej (symptom osiągnięcia zbieżności).

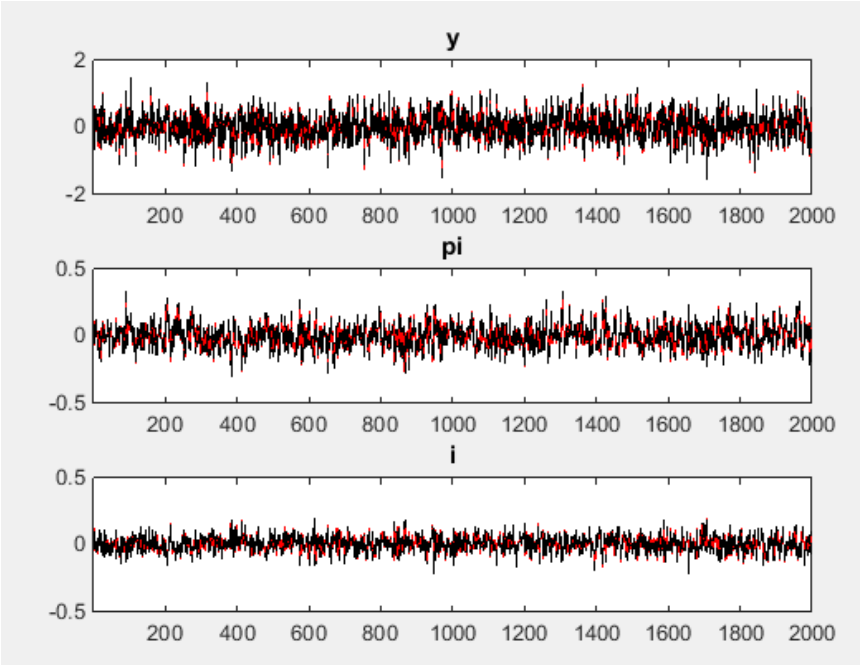


Jak ocenić wyniki estymacji? (3a)

- Obejrzyjmy również wnikliwie ścieżki zmiennych obserwowalnych (nasze dane) vs wygładzonych (tzn. pozbawionych błędów pomiaru).
  - Gdy zmienne wygładzone różnią się bardzo, np. usunięta jest z nich większość wariancji, mogliśmy osiągnąć efekt odwrotny do zamierzonego.
  - Wariancja zawarta w tych zmiennych nie pasowała do modelu, a więc została ona "wypchnięta" poza niego i wykorzystane dane w zasadzie w ogóle nie wzięły udziału w ustalaniu wartości parametrów strukturalnych.
  - Innymi słowy, model obraził się na rzeczywistość, że do niego nie przystaje, i "się w sobie zamknął"... (tzn. przypisał lwią część wariancji zmiennych błędom pomiaru – empiryczna porażka modelu!).

Jak ocenić wyniki estymacji? (3b)

Czerwona linia (bez błędu pomiaru) powinna mieć gładszy przebieg niż czarna (z błędem pomiaru). Jednak nie powinna być zupełnie płaska w porównaniu z czarną, bo to sugeruje kłopot z dopasowaniem.



Plan prezentacji

Zmienne w modelu DSGE (1)

Zanim przystąpimy do estymacji, przypomnijmy sobie, że...

- wiele zmiennych w modelu (y, c) to procentowe odchylenia od stanu ustalonego (gdzie ich szukać?)
- inflacja (π) mierzona jest wokół stanu ustalonego w którym ceny nie zmieniają się (a cel inflacyjny NBP?)
- produkcja i konsumpcja (y, c) mierzą ilości, a więc interesują nas kategorie realne (*volumes*), a nie nominalne
- stopy procentowe (i) winny być wyrażone w rzędzie wielkości analogicznym do procentowych odchyleń od stanu ustalonego, tzn. konsekwentnie 1 albo 0.01 dla stopy procentowej równej 1% i jednoprocentowego odchylenia od stanu ustalonego dla produkcji
- nie ma czegoś takiego jak dane o sektorze T i NT, musimy je sobie sami skonstruować...
- w modelu nie mamy wbudowanych żadnych mechanizmów sezonowości (np. egzogeniczna pogoda czy schematy zatrudnienia), dlatego korzystamy z danych odsezonowanych przez Eurostat

Zmienne w modelu DSGE (2)

A zatem będziemy musieli wykonać kilka operacji:

- ściągnięcie odpowiednich szeregów (zob. **dane.xls**, zakładka **Eurostat**)
- kalibracja parametrów, których nie będziemy estymować (zob. **dane.xls**, 2. zakładka – kolumny M i N, 3. zakładka – kolumny B, C i D, 4. zakładka)
- konstrukcja zmiennych sektorowych (y, π) na podstawie danych o wartości dodanej i jej deflatorze (**dane.xls**, zakładka **dane_do_matlaba** i poprzednie)
- logarytmowanie indeksów ilości (y, c) oraz cen i kursu walutowego, wyrażenie tych ostatnich jako logarytmicznych przyrostów (zakładka **dane_do_matlaba**)
- odjęcie średniej od wszystkich szeregów (zob. **import_danych_matlab.m**)

Zmienne w modelu DSGE (3)

- wyrażenie y oraz c jako odchyleń od stanu ustalonego poprzez logarytmiczne odchylenia od trendu wyznaczonego filtrem Hodricka-Prescotta
  - w Matlabie jest polecenie hpfilter w **Econometrics Toolbox**
  - jeżeli nie mamy tego dodatku lub mamy Octave, korzystamy z dołączonego pliku **hpfilter.m**
  - filtrację można też przeprowadzić łatwo w innym programie, np. w Gretlu
- w naszym przykładzie filtrujemy również stopę nominalną w Polsce ze względu na obniżający się cel inflacyjny w początkowym okresie próby

Ćwiczenie 1

Zanim przejdziemy do prawdziwej konfrontacji z danymi (która może rozczarować), bezpiecznym ćwiczeniem będzie:

1. symulacja stochastyczna modelu gospodarki zamkniętej;
2. zapisanie wygenerowanych ścieżek;
3. estymacja parametrów modelu na podstawie tych ścieżek.

Całość zapiszemy jednym pliku modelu (mod).
Uwaga! Potrzebne będzie polecenie zapisujące wygenerowane przez Dynare zmienne do pliku zewnętrznego Octave/Matlaba (jako wektory):

```
dynasave('plik_danych.mat') y i pi;
```

Ćwiczenie 2

Przeprowadź estymację modelu z pliku **NK_open_estimation.mod** dla strefy euro oraz kraju innego niż Polska, np. Czech. Samodzielnie skonstruuj zbiór danych i dokonaj odpowiednich korekt.

Ćwiczenie 3

Przeprowadź estymację dla modelu gospodarki otwartej należącego do strefy euro od początku, np. Irlandii. Przyjmij w uproszczeniu, że dane dla całej strefy euro nie obejmują tego kraju (lub, innymi słowy, że to kraj na tyle mały, by nie ważyć na nich istotnie). Jakich zmian należy dokonać w modelu i zbiorze danych? (przypomnijmy sobie wynik jednego z zadań z poprzedniego tematu)