Bayesiánská analýza

VI. Lineární regresní model s obecnou kovarianční maticí

Obsah tématu

Obsah tématu

Úvod

- Šestá kapitola z Koop (2003) resp. učebního textu.
- Heteroskedasticita náhodných složek.
- Autokorelace náhodných složek.
- Hierarchická apriorní hustota.
- Modely zdánlivě nesouvisejících regresí.

Předpoklady

- Model:
$$y = X\beta + \epsilon.$$
- Předpoklady:
  1. $\epsilon \sim N(0_N, h^{-1}\Omega)$.
  2. Všechny prvky X jsou buď nenáhodné veličiny nebo nezávislé náhodné veličiny vzhledem k ϵ.
- Různé modely podle podoby Ω.

Obsah tématu

Transformace modelu

- Existence matice P: $P\Omega P' = I_N$.
- Transformovaný model:

$$y^* = X^* \beta + \epsilon^*,$$

  kde $y^* = Py$, $X^* = PX$ a $\epsilon^* = P\epsilon$.
- $\epsilon^* \sim N(0_N, h^{-1} I_N)$.

Implikace

- Pokud Ω známa $\rightarrow$ transformace dat (standardní bayesiánská analýza).
- Pokud Ω neznámá $\rightarrow$ podmíněno Ω posteriorní hustoty pro β a h standardní podobu + $p(\Omega|y, \beta, h)$ závisí na přesné formě Ω.
- Ostatní podmíněné hustoty stejné jako dříve (podmíněné Ω).

Věrohodnostní funkce

- Z vlastností vícerozměrného normálního rozdělení:

$$p(y|\beta, h, \Omega) = \frac{h^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} |\Omega|^{-\frac{1}{2}} \left\{ \exp\left[-\frac{h}{2}(y - X\beta)'\Omega^{-1}(y - X\beta) \right] \right\}.$$

- S transformovanými daty:

$$p(y^*|\beta, h, \Omega) = \frac{h^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \left\{ \exp\left[-\frac{h}{2}(y^* - X^*\beta)'(y^* - X^*\beta) \right] \right\}.$$

Zobecněná metoda nejmenších čtverců

$$
\begin{aligned}
\nu &= N - k. \\
\widehat{\beta}(\Omega) &= (X^{*\prime}X^{*})^{-1}X^{*\prime}y^{*} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y, \\
s^2(\Omega) &= \frac{(y^{*} - X^{*}\widehat{\beta}(\Omega))'(y^{*} - X^{*}\widehat{\beta}(\Omega))}{\nu} \\
&= \frac{(y - X\widehat{\beta}(\Omega))'\Omega^{-1}(y - X\widehat{\beta}(\Omega))}{\nu}, \\
p(y|\beta, h, \Omega) &= \frac{1}{(2\pi)^{\frac{N}{2}}} \\
&\quad \times \left\{ h^{\frac{1}{2}} \exp\left[-\frac{h}{2}(\beta - \widehat{\beta}(\Omega))'X'\Omega^{-1}X(\beta - \widehat{\beta}(\Omega)) \right] \right\} \\
&\quad \times \left\{ h^{\frac{\nu}{2}} \exp\left[-\frac{h\nu}{2s^{-2}(\Omega)} \right] \right\}.
\end{aligned}
$$

Apriorní hustota

- Normální-gama apriorní hustotu pro β a h a obecné značení $p(\Omega)$ pro Ω:

$$p(\beta, h, \Omega) = p(\beta)p(h)p(\Omega),$$

kde

$$p(\beta) = f_N(\beta|\underline{\beta}, \underline{V}),$$
$$p(h) = f_G(h|\underline{\nu}, \underline{s}^{-2}).$$

Posteriorní hustota

- Sdružená hustota nemá tvar známé hustoty:

$$
\begin{aligned}
p(\beta, h, \Omega|y) \propto\ & p(\Omega) \\
& \times \left\{ \exp\left[-\frac{1}{2} \left\{ h(y^* - X^*\beta)'(y^* - X^*\beta) \right.\right.\right. \\
& \left.\left.\left. + (\beta - \underline{\beta})' \underline{V}^{-1} (\beta - \underline{\beta}) \right\} \right] \right\} \\
& \times h^{\frac{N+\underline{\nu}-2}{2}} \exp\left[-\frac{h\underline{\nu}}{2\underline{s}^{-2}} \right].
\end{aligned}
$$

- Odvození podmíněných hustot.

Podmíněná posteriorní hustota pro β

- Vícerozměrné normální rozdělení:

$$\beta | y, h, \Omega \sim N(\overline{\beta}, \overline{V}),$$

kde

$$\overline{V} = (\underline{V}^{-1} + hX'\Omega^{-1}X)^{-1},$$
$$\overline{\beta} = \overline{V}(\underline{V}^{-1}\underline{\beta} + hX'\Omega^{-1}X\widehat{\beta}(\Omega)).$$

Podmíněná posteriorní hustota pro h

- Gama rozdělení:

$$h|y,\beta \sim G(\overline{s}^{-2}, \overline{\nu}),$$

kde

$$\overline{\nu} = N + \underline{\nu},$$
$$\overline{s}^2 = \frac{(y - X\beta)'\Omega^{-1}(y - X\beta) + \underline{\nu s}^2}{\overline{\nu}}.$$

Podmíněná posteriorní hustota pro Ω

- Jádrová hustota:

$$p(\Omega|y,\beta,h) \propto p(\Omega)|\Omega|^{-\frac{1}{2}}\left\{\exp\left[-\frac{h}{2}(y-X\beta)'\Omega^{-1}(y-X\beta)\right]\right\}$$

- Obecně nenabývá podoby známého rozdělení.
- Následně konkretizace + odvození posteriorních simulátorů.
- Pokud jsme schopni generovat výběry z $p(\Omega|y,\beta,h) \rightarrow$ Gibbsův vzorkovač.

Obsah tématu

Úvod

- $var(\epsilon_i) = h^{-1}\omega_i$ pro $i = 1, \ldots, N$:

$$\Omega = \begin{bmatrix} \omega_1 & 0 & \cdot & \cdot & 0 \\ 0 & \omega_2 & 0 & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & \omega_N \end{bmatrix}.$$

- Obecně předpokládáme: $\omega_i = h(z_i, \alpha)$, kde $h()$ je kladná funkce a z_i vektor všech nebo některých vysvětlujících proměnných.
- Obvyklá volba:

$$h(z_i, \alpha) = (1 + \alpha_1 z_{i1} + \alpha_2 z_{i2} + \ldots + \alpha_p z_{ip})^2.$$

Princip

- Jsme schopni vyjádřit výraz pro Ω.
- Ω závisí na parametrech α.
- Vhodný simulátor: Metropolis-within-Gibbs.
- $p(\beta|y,h,\alpha)$ normální; $p(h|y,\beta,\alpha)$ gama; potřeba generování vzorků z $p(\alpha|y,\beta,h)$.

$$p(\alpha|y,\beta,h) \propto p(\alpha)|\Omega(z_i,\alpha)|^{-\frac{1}{2}} \times \left\{\exp\left[-\frac{h}{2}(y-X\beta)'\Omega(z_i,\alpha)^{-1}(y-X\beta)\right]\right\}.$$

- Např. Random Walk Chain Metropolis-Hastings algoritmus + Bayesův faktor pro např. $\alpha_1 = \ldots = \alpha_p = 0$ (přístup Gelfanda-Deye) + HPDI, p-hodnoty.

Empirická ilustrace

- Viz Koop (2003) – bude doplněno.

Obsah tématu

Obecné principy

- Platí:

$$\Omega = \begin{bmatrix} \omega_1 & 0 & \cdot & \cdot & 0 \\ 0 & \omega_2 & 0 & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & \omega_N \end{bmatrix}.$$

- Neznáme, neumíme nebo nechceme specifikovat podobu pro ω_i.
- Jak odhadnout na základě N pozorování $N + k + 1$ parametrů β, h a $\omega = (\omega_1, \dots, \omega_N)'$?

Užitečnost postupu

1. Využití tzv. *hierarchických priorů* – způsob tvorby více flexibilních na parametry bohatých modelů pro statistickou analýzu.
2. Zavedení konceptu vztahujícího se k flexibilnímu ekonometrickému modelování (umožnění uvolnění předpokladu normality náhodné složky).

Značení

- Apriorní hustota $p(\omega)$ pro N-rozměrný vektor parametrů $\omega \rightarrow$ přesnost chyby:

$$\lambda \equiv (\lambda_1, \lambda_2, \ldots, \lambda_N)' \equiv (\omega_1^{-1}, \omega_2^{-1}, \ldots, \omega_N^{-1})'.$$

- Předpokládejme apriorní hustotu pro λ:

$$p(\lambda) = \prod_{t=1}^{N} f_G(\lambda_i|1, \nu_\lambda).$$

- Závisí na hyperparametru ν_λ, který si volíme.
- Předpokládá se, že každé λ_i pochází z téhož rozdělení (i.i.d. výběry z gama rozdělení).
- Vypořádání se s problémy vysoké dimenzionality vektoru $\lambda \rightarrow$ rozptyly náhodných složek se budou vzájemně lišit, ale budou pocházet ze stejného rozdělení $\rightarrow$ flexibilní model s dostatečně pevnou strukturou pro možnou statistickou analýzu.

Důvod volby

- Proč gama rozdělení se střední hodnotou 1.0?
- Model stejný jako lineární regresní model s náhodnými složkami z i.i.d. Studentova t-rozdělení s ν_λ stupni volnosti:

$$p(\epsilon_i) = f_t(\epsilon_i | 0, h^{-1}, \nu_\lambda).$$

- Studentovo t-rozdělení podobné normálnímu rozdělení $\times$ tlustší konce a je více flexibilnější.
- Výhoda: rámec normálního lineárního regresního modelu (využití výpočetních postupů pro posteriorní simulátor pro LRM s nezávislými t-rozdělenými chybami).

Obecné vysvětlení

- Mix specifických normálních rozdělení $\rightarrow$ tvorba mnohem flexibilnějšího rozdělení váženým průměrem více normálních rozdělení.
- Kompozice (mixture) normálních modelů = mocný nástroj, kdy ekonomická teorie nenabízí specifikaci věrohodnostní funkce a my chceme být dostatečně flexibilní.
- Naše chápání heteroskedasticity v neznámé podobě = poměrný mix normálních rozdělení (*scale mixture of Normals*).
- Předpoklad $\epsilon_i \sim N(0, h^{-1}\lambda_i^{-1})$ s naší apriorní hustotou pro λ_i = rozdělení chyb je váženým průměrem (mixem) různých normálních rozdělelní s různými rozptyly (tj. různá měřítka – *scales*) a stejné střední hodnoty (tj. všechny chyby mají nulovou střední hodnotu).
- Pokud kompozice vytvořena za použití gama hustot $f_G(\lambda_i|1, \nu_\lambda) \rightarrow$ t-rozdělení.
- Jiné hustoty než $f_G(\lambda_i|1, \nu_\lambda) \rightarrow$ jiná (flexibilnější) rozdělení.

Další uvolnění

- Pokud ν_λ neznámé: neznámý parametr.
- V bayesovském konceptu potřeba apriorní hustoty: $p(\nu_\lambda)$.
- Specifikace apriorní hustoty pro λ ve dvou krocích:

$$p(\lambda) = \prod_{t=1}^{N} f_G(\lambda_i | 1, \nu_\lambda) \quad \text{a} \quad p(\nu_\lambda).$$

- Alternativně: $p(\lambda|\nu_\lambda)p(\nu_\lambda)$.
- *Hierarchické apriorní hustoty* = apriorní hustoty zapsána ve dvou či více krocích.
- Není nutné $\rightarrow$ hierarchický prior lze zapsat v nehierarchickém pojetí.
- V našem případě: $p(\lambda) = \int p(\lambda|\nu_\lambda)p(\nu_\lambda)d\nu_\lambda$.

Problémy

- Pro posteriorní hustotu nemusí existovat střední hodnoty a směrodatné odchylky.
- Geweke – pokud použijeme neinformativní apriorní hustotu pro β ($p(\beta) \propto 1$ na intervalu $(-\infty, \infty)$):
  1. Neexistuje posteriorní střední hodnota, pokud $p(\nu_\lambda)$ není nulová na intervalu $(0, 2\rangle$.
  2. Neexistuje posteriorní směrodatná odchylka, pokud $p(\nu_\lambda)$ není nulová na intervalu $(0, 4\rangle$.
- Neinformativní apriorní hustotu pro ν_λ:

$$p(\nu_\lambda) \propto 1 \quad \nu_\lambda(0, \infty).$$

- $\frac{p(\nu_\lambda \leq 100)}{p(\nu_\lambda > 100)} = 0 \Rightarrow$ hodně informativní (náhodné složky normálně rozděleny).

Gibbsův vzorkovač

- Formálně plně podmíněné hustoty pravděpodobnosti: $p(\beta|y, h, \lambda, \nu_\lambda)$ a $p(h|y, \beta, \lambda, \nu_\lambda)$ $\times$ za podmínky λ nám ν_λ nepřináší žádnou novou informaci.
- $p(\beta|y, h, \lambda, \nu_\lambda) = p(\beta|y, h, \lambda)$ a $p(h|y, \beta, \lambda, \nu_\lambda) = p(h|y, \beta, \lambda)$ již odvozeny.
- Odvození podmíněné hustoty pro λ zřejmé: dosazení apriorní hustoty danou do obecného tvaru podmíněné posteriorní hustoty pro Ω (jedná se o funkci proměnných λ_i).
- Zjišťujeme, že λ_i jsou navzájem nezávislé:

$$p(\lambda|y, \beta, h, \nu_\lambda) = \prod_{i=1}^{N} p(\lambda_i|y, \beta, h, \nu_\lambda),$$

$$p(\lambda_i|y, \beta, h, \nu_\lambda) = f_G\left(\lambda_i|\frac{\nu_\lambda + 1}{h\epsilon_i^2 + \nu_\lambda}, \nu_\lambda + 1\right).$$

Výpočet pro ν_λ

- Specifikace apriorní hustoty: exponenciální rozdělení

$$p(\nu_\lambda) = f_G(\nu_\lambda | \underline{\nu}_\lambda, 2)$$

- Jiné apriorní hustoty = drobné úpravy algoritmu posteriorní simulace.
- $p(\nu_\lambda | y, \beta, h, \lambda)$ snadné odvodit (ν_λ nevstupuje do věrohodnostní funkce) $\Rightarrow p(\nu_\lambda | y, \beta, h, \lambda) = p(\nu_\lambda | \lambda)$.
- Z Bayesova teorému: $p(\nu_\lambda | \lambda) \propto p(\lambda | \nu_\lambda) p(\nu_\lambda)$.
- Jádrová podmíněná posteriorní hustota:

$$p(\nu_\lambda | y, \beta, h, \lambda) \propto \left(\frac{\nu_\lambda}{2}\right)^{\frac{N\nu_\lambda}{2}} \Gamma\left(\frac{\nu_\lambda}{2}\right)^{-N} \exp(-\eta \nu_\lambda)$$

kde

$$\eta = \frac{1}{\underline{\nu}_\lambda} + \frac{1}{2} \sum_{i=1}^{N} [\ln(\lambda_i^{-1}) + \lambda_i].$$

- Nestandardní hustota → MH algoritmus (doporučen i jiný algoritmus – *acceptance sampling*).

Další analýza

- Pro řadu hypotéz (např. $\beta_j = 0$) – Savage-Dickey density ratio.
- Test náznaku odchýlení náhodných složek od normality: porovnání $M_1 : \nu_\lambda \to \infty$ a $M_2 : \nu_\lambda$ jako konečné číslo.
- Bayesův faktor za použití metody Gelfanda-Deye = posteriorní simulátor pro každý z těchto modelů.
- Alternativně lze využít predikční p-hodnoty nebo HPDI.
- Predikční analýza probíhá standardním způsobem.

Empirická ilustrace

- Viz Koop (2003) – bude doplněno.

Obsah tématu

Úvod

- Korelace proměnných v čase (setrvačnost v preferencích, proces přizpůsobení).
- Náhodná složka jižne $N(0_N, h^{-1}I_N)$.
- y_t pro $t = 1\dots, T$.
- Příklad *autoregresního procesu řádu 1* $(AR(1))$:

$$\epsilon_t = \rho\epsilon_{t-1} + u_t,$$

kde u_t je i.i.d. $N(0, h^{-1})$.

Základní pojmy a vlastnosti

- Obecně řada z_t od $-\infty$ do ∞.
- Pozorujeme pro $t = 1, \ldots, T$.
- *Kovariančně stacionární (či slabě stacionární)*, pokud pro každé t a s platí:

$$\begin{aligned} E(z_t) &= E(z_{t-s}) = \mu, \\ var(z_t) &= var(z_{t-s}) = \gamma_0, \\ cov(z_t, z_{t-s}) &= \gamma_s, \end{aligned}$$

  kde μ, γ_0 a γ_s jsou konečné hodnoty.
- Diference m-tého řádu pro $m > 1$: $\Delta^m z_t = \Delta^{m-1} z_t - \Delta^{m-1} z_{t-1}$.
- γ_s = *autokovariační funkce* $\rightarrow$ *autokorelační funkce* $= \frac{\gamma_s}{\gamma_0}$ pro $s = 0, \ldots, \infty$.

$AR(1)$ proces

- ϵ_t jako funkce u_{t-s} pro $s = 0, \ldots, \infty$:

$$\epsilon_t = \sum_{s=0}^{\infty} \rho^s u_{t-s}.$$

- Problém při výpočtu střední hodnoty, rozptylu a kovariance $\epsilon_t \to \rho^s$ bude nekonečno pro $|\rho| > 1$ (+ pro $\rho = 1$).
- Podmínka stacionarity: $|\rho| < 1$.
- $|\rho| < 1 \Rightarrow E(\epsilon_t) = 0$,

$$\gamma_0 = var(\epsilon_t) = h^{-1} \sum_{s=0}^{\infty} \rho^{2s} = \frac{1}{h(1-\rho^2)},$$

$$\gamma_s = cov(\epsilon_t, \epsilon_{t-s}) = \frac{\rho^s}{h(1-\rho^2)}.$$

- Autokovarianční funkce γ_s klesá s rostoucím s.

Zápis Ω

- Kovarianční matice pro ϵ: $h^{-1}\Omega$.

$$\Omega = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdot & \rho^{T-1} \\ \rho & 1 & \rho & \cdot & \cdot \\ \rho^2 & \rho & \cdot & \cdot & \rho^2 \\ \cdot & \cdot & \cdot & \cdot & \rho \\ \rho^{T-1} & \cdot & \rho^2 & \rho & 1 \end{bmatrix}$$

Zobecnění

- $AR(p)$ proces:

$$\epsilon_t = \rho_1 \epsilon_{t-1} + \ldots + \rho_p \epsilon_{t-p} + u_t.$$

- Výpočet střední hodnoty, rozptylu a autokovarianční funkce (pro bayesovskou analýzu není potřeba).
- *Operátor zpoždění*: L.
- $L\epsilon_t = \epsilon_{t-1}$ resp. $L^m \epsilon_t = \epsilon_{t-m}$.
- $AR(p)$ proces:

$$(1 - \rho_1 L - \ldots - \rho_p L^p)\epsilon_t = u_t \quad \text{nebo} \quad \rho(L)\epsilon_t = u_t.$$

- $\rho(L) = (1 - \rho_1 L - \ldots - \rho_p L^p)$: polynom řádu p pro operátor zpoždění.
- $AR(p)$ je stacionární $\Leftrightarrow$ kořeny rovnice $\rho(z) = 0$ jsou všechny v absolutní hodnotě větší než jedna.
- $\rho = (\rho_1, \ldots, \rho_p)'$ a Φ je stacionární oblast modelu.

Možnosti transformace

- Posteriorní simulátro z předchozích vztahů pro obecnou matici Ω.
- Specifická transformace $\rightarrow$ vztahy v jednoduché podobě.
- Specifikujeme Ω za předpokladu $AR(p)$ procesu chyb.
- Odvození matice P: $P\Omega P' = I$.
- Transformace.

Alternativní postup transformace

- Původní model:
$$y_t = x_t'\beta + \epsilon_t,$$
kde $x_t = (1, x_{t2}, \dots, x_{tk})'$.
- Přenásobení obou stran pomocí $\rho(L)$.
- $y_t^* = \rho(L)y_t$ a $x_t^* = \rho(L)x_t$:
$$y_t^* = x_t^{*\prime}\beta + u_t.$$
- u_t je i.i.d. $N(0, h^{-1})$.

„Problém"

- Transformované hodnoty pro $t \leq p$?
- Příklad: y_1^* závisí na $y_0, \ldots, y_{1-p}$.
- Ošetření počátečních podmínek $\rightarrow$ problém nestacionarity AR procesu (blízkosti nestacionaritě).
- Obvykle: práce s věrohodnostní funkcí založenou na datech od $t = p + 1, \ldots, T$.
- Pokud p relativně malé vzhledem k T $\rightarrow$ výsledná aproximace dobrá.
- y_t^* a x_t^* pro $t = p + 1, \ldots, T$ nezávisí na nepozorovaných zpožděných hodnotách.

Značení

- Žádné speciální značení.
- Vycházíme jen z dat od $t = p + 1, \ldots, T$.
- y, y^*, ϵ a ϵ^*: vektory rozměru $T - p$.
- Matice X a X^*: rozměr $(T - p) \times k$.
- Gibbsův vzorkovač s využitím výsledků předchozích částí: $p(\beta|y, h, \rho)$ a $p(h|y, \beta, \rho)$ jsou dány dříve.
- $p(\rho|y, \beta, h)$ lze odvodit: za podmínky β a h je ϵ_t pro $t = p + 1, \ldots, T$ známé a $AR(p)$ proces je NLRM (se známým rozptylem chyb) a s koeficienty danými vektorem ρ.

Podmíněná posteriorní hustota pro β

- Nezávislá normální-gama apriorní hustota pro β a h.

$$\beta|y, h, \rho \sim N(\overline{\beta}, \overline{V}),$$

kde

$$\overline{V} = (\underline{V}^{-1} + hX^{*\prime}X^{*})^{-1},$$
$$\overline{\beta} = \overline{V}(\underline{V}^{-1}\underline{\beta} + hX^{*\prime}y^{*}).$$

Podmíněná posteriorní hustota pro h

- Gama rozdělení:

$$h|y, \beta\rho \sim G(\overline{s}^{-2}, \overline{\nu}),$$

kde

$$\overline{\nu} = T - p + \underline{\nu},$$
$$\overline{s}^2 = \frac{(y^* - X^*\beta)'(y^* - X^*\beta) + \underline{\nu s}^2}{\overline{\nu}}.$$

Hustoty pro ρ

- Posteriorní hustota pro ρ závisí na své apriorní hustotě.
- Předpokládáme vícerozměrné normální rozdělelní omezené na stacionární oblast:

$$p(\rho) \propto f_N(\rho|\underline{\rho}, \underline{V}_\rho)1(\rho \in \Phi),$$

- Podmíněná posteriorní hustota:

$$p(\rho|y, \beta, h) \propto f_N(\rho|\overline{\rho}, \overline{V}_\rho)1(\rho \in \Phi),$$

kde

$$\overline{V}_\rho = (\underline{V}_\rho^{-1} + hE'E)^{-1},$$
$$\overline{\rho} = \overline{V}_\rho(\underline{V}_\rho^{-1}\underline{\rho} + hE'\epsilon).$$

- E je matice $(T - p) \times p$ s t-tým řádkem $\epsilon_{t-1}, \ldots, \epsilon_{t-p}$.

Komplikace

- Gibbsův vzorkovač: sekvenční výběry z podmíněných hustot.
- Podmíněná posteriorní hustota pro $\rho =$ omezené vícerozměrné normální rozdělení $\rightarrow$ drobná komplikace.
- Výběry z neomezeného rozdělení a vypuštění vzorků mimo stacionární oblast (pokud $\overline{\rho}$ leží uvnitř této oblasti nebo alespoň nepříliš daleko od ní).
- Alternativně lze odvodit Metropolis-Hastings algoritmus.

Další analýza

- Predikční analýza standardním způsobem; predikční p-hodnoty nebo HPDI; Bayesův faktor pro ověření použitím Savage-Dickeyeho poměru hustot nebo za pomocí metody Gelfanda-Deye.
- Savage-Dickey density ratio: úplné funkce hustoty pravděpodobnosti $p(\rho|y, \beta, h)$ a $p(\rho|y)$.
- Pro $p = 1$ není problém ($p(\rho|y, \beta, h)$ omezené jednorozměrné normální).
- Pro $p > 1$ stacionární oblast nelineární (analytické vyjádření $p(\rho|y, \beta, h)$ obtížné).
- Aproximativní výpočet integrační konstanty:

$$p(\rho|y, \beta, h) = \frac{f_N(\rho|\overline{\rho}, \overline{V}_\rho) 1(\rho \in \Phi)}{\int_\Phi f_N(\rho|\overline{\rho}, \overline{V}_\rho) d\rho}.$$

- Vzorky z $f_N(\rho|\overline{\rho}, \overline{V}_\rho)$ a vyhození výběrů mimo stacionární oblast.
- $\int_\Phi f_N(\rho|\overline{\rho}, \overline{V}_\rho) d\rho$: podíl vzorků (z celkového počtu), který nám zůstane $\rightarrow 1 - \int_\Phi f_N(\rho|\overline{\rho}, \overline{V}_\rho) \mathrm{d}\rho$.

Empirická ilustrace

- Viz Koop (2003) – bude doplněno.

Obsah tématu

Motivace

- „Seemingly Unrelated Regressions“ (SUR) model.
- Analýza spotřeby (poptávka po různých kategoriích), poptávka po výrobních faktorech (pro každý z faktorů).
- Klasické ekonometrie: redukovaný tvar modelu simultánních rovnic, VAR model.
- Někdy práce s jednotlivými rovnicemi dostatečná $\times$ SUR model pro „lepší“ odhad.

Zápis

- SUR model:
$$y_{mi} = {x_{mi}}'\beta_m + \epsilon_{mi}.$$
- $i = 1, \ldots, N$: N pozorování pro $m = 1, \ldots, M$ rovnic.
- y_{mi}: i-té pozorování závisle proměnné v rovnici m.
- x_{mi}: k_m-rozměrný vektor obsahující i-té pozorování vektoru vysvětlujících proměnných v m-té rovnici.
- β_m: k_m-rozměrný vektor regresních koeficientů pro m-tou rovnici.
- Vysvětlující proměnné mohou být různé v jendotlivých rovnicích.

Maticové vyjádření

- Přepsání do vektorů a matic:

$$y_i = \begin{pmatrix} y_{1i} \\ \cdot \\ \cdot \\ y_{Mi} \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \cdot \\ \cdot \\ \beta_M \end{pmatrix} \quad X_i = \begin{pmatrix} x'_{1i} & 0 & \cdot & \cdot & 0 \\ 0 & {x_{2i}}' & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & x'_{Mi} \end{pmatrix}.$$

- $k = \sum_{m=1}^{M} k_m \rightarrow y_i = X_i\beta + \epsilon_i$.

Maticové vyjádření II

- Pozorování dohromady:

$$y = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_N \end{pmatrix} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \cdot \\ \cdot \\ \epsilon_N \end{pmatrix} \quad X = \begin{pmatrix} X_1 \\ \cdot \\ \cdot \\ X_N \end{pmatrix}.$$

- Zápis: $y = X\beta + \epsilon$.
- SUR model – lineární regresní model.

Předpoklad

- NLRM: ϵ_{mi} je i.i.d. $N(0, h^{-1})$ pro všechna i a m.
- Předpoklad: ϵ_i je i.i.d. $N(0, H^{-1})$ pro $i = 1, \ldots, N$, kde H je matice přesností chyb rozměru $M \times M$.
- $\epsilon \sim N(0, \Omega)$, kde Ω je blokově-diagonální matice rozměru $NM \times NM$:

$$\Omega = \begin{pmatrix} H^{-1} & 0 & \cdot & \cdot & 0 \\ 0 & H^{-1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & H^{-1} \end{pmatrix}.$$

- Nevystupuje zde h (žádný rozdíl oproti předchozímu, možno dodatečně přidat).

Apriorní hustota

- Rozšíření nezávislého normálního-gama rozdělení do podoby nezávislého normálního-Wishartova rozdělení:

$$p(\beta, H) = p(\beta)p(H),$$

kde

$$p(\beta) = f_N(\beta|\underline{\beta}, \underline{V}),$$
$$p(H) = f_W(H|\underline{\nu}, \underline{H}).$$

- Wishartovo rozdělení: maticové zobecnění gama rozdělení.
- $E(H) = \underline{\nu H}$ a neinformativnost pro volbu $\underline{\nu} = 0$ a $\underline{H}^{-1} = 0_{M\times M}$.
- Možné jiné apriorní hustoty (přirozeně konjugované mnohdy restriktivní $\times$ analytické výsledky).

Posteriorní hustota β

- Gibbsův vzorkovač na základě známých vztahů: inverze $NM \times NM$-rozměrné matice Ω.
- Bloková struktura $\rightarrow$ částečně analytická inverze $\rightarrow$ obvyklá podoba $p(\beta|y, H)$ a $p(H|y, \beta)$:

$$\beta|y, H \sim N(\overline{\beta}, \overline{V})$$

kde

$$\overline{V} = \left(\underline{V}^{-1} + \sum_{i=1}^{N} X_i{}' H X_i \right)^{-1},$$

$$\overline{\beta} = \overline{V} \left(\underline{V}^{-1} \underline{\beta} + \sum_{i=1}^{N} X_i{}' H y_i \right).$$

Posteriorní hustota H

- Podmíněná hustota odpovídá Wishartovu rozdělení:

$$H|y,\beta \sim W(\overline{\nu}, \overline{H})$$

kde

$$\overline{\nu} = N + \underline{\nu},$$
$$\overline{H} = \left[\underline{H}^{-1} + \sum_{i=1}^{N}(y_i - X_i\beta)(y_i - X_i\beta)'\right]^{-1}.$$

- Generátory náhodných čísel z Wishartova rozdělení jsou k dispozici $\Rightarrow$ snadná implementace Gibbsova vzorkovače.

Závěr

- Standardně predikční analýza.
- Ověření kvality modelu pomocí predikční p-hodnoty a HPDI.
- Posteriorní podíl šancí pomocí Savageho-Dickeyho poměru hustot.

Empirická ilustrace

- Viz Koop (2003) – bude doplněno.