

Застосування методу Монте Карло для марковських ланцюгів до оцінювання моделі стохастичної волатильності

В моделі такого типу мінлива структура дисперсії описується стохастичним, наприклад, нормальним розподіленим процесом. Приклад оцінювання такої моделі методом квазі-максимальної правдоподібності, наведено у роботі[1]. В роботі [2] досліджено поведінку методу квазі-максимальної правдоподібності для оцінювання параметрів гетероскедастичних процесів та в тому числі моделі стохастичної волатильності на різних вибірках. Зроблені висновки стосовно того, що метод не дуже якісно оцінює $\ln[h(k)]$, даючи зміщені оцінки. Існують також проблеми стосовно логарифмічного перетворення, яке неможливо виконати при нульових чи близьких до нуля значеннях. Одним із обмежуючих факторів є те, що не кожна модель гетероскедастичного процесу (зокрема, моделі стохастичної волатильності) може бути представлена у вигляді, лінійному стосовно $h(k)$, а саме це використовується у згаданому методі.

Структура стохастичної моделі волатильності може бути представлена таким чином:

$$\begin{aligned}y(k) &= a_0 + a_1 y(k-1) + \varepsilon(k), \\ \varepsilon(k) &= \sqrt{h(k)} v(k), \quad \{v(k)\} \sim N(0,1), \\ \ln[h(k+1)] &= \alpha_0 + \alpha_1 \ln[h(k)] + \eta(k), \\ \{\eta(k)\} &\sim N(0, \sigma_\eta^2),\end{aligned}$$

де $y(k)$ – основна змінна; $\varepsilon(k)$, $\eta(k)$ – незалежні стохастичні процеси білого шуму; $h(k)$ – умовна дисперсія; a_0 , a_1 , α_0 , α_1 – коефіцієнти моделі. Процес авторегресії першого порядку з інновацією $\varepsilon(k)$ пояснює можливе існування авторегресії у процесі $\{y(k)\}$. Для оцінювання таких моделей

використовують метод квазі-максимальної правдоподібності та байєсівський підхід у поєднанні з методом Монте Карло.

Постановка задачі. Розробити альтернативний підхід до розв'язання задачі оцінювання вектора параметрів нелінійної моделі для прогнозування волатильності гетероскедастичних процесів довільної природи типу $\{y(k)\} \sim N[\bar{y}, h(k)]$, $h(k) \neq const$ (тут $y(k)$ – досліджуваний процес; \bar{y} – середнє значення і $h(k)$ – умовна дисперсія процесу). Створити теоретичне підґрунтя для використання цього підходу, розглянути приклади оцінювання моделі стохастичної волатильності.

Байєсівський підхід до оцінювання параметрів моделей. Основним елементом байєсівського підходу є теорема Байєса (ТБ). Загалом ТБ стосується обчислення умовних ймовірностей подій загального характеру. Нехай $\mathbf{A} = [A_1, \dots, A_n]$ – множина взаємовиключних подій, які повністю описують деякий процес. Розглянемо деяку подію A_j , виникнення якої зв'язане з подією B . Теорема Байєса дає можливість визначити умовну ймовірність настання події A_j за умови, що мала місце подія B , використовуючи при цьому інформацію стосовно умовної ймовірності настання події B (за умови, що мала місце подія A_j). Виходячи з наведеного формульовання призначення теореми Байєса, можна сказати, що вона стосується так званої “інверсної ймовірності”. Теорема Байєса для довільних подій записується так:

$$p\{A_j | B\} = \frac{p\{B | A_j\} p\{A_j\}}{\sum_{k=1}^n p\{B | A_k\} p\{A_k\}}, \quad p\{B\} \neq 0. \quad (2.4.1)$$

Доводиться вона через формулу повної ймовірності:

$$p\{B\} = \sum_{j=1}^n p\{B | A_j\} p\{A_j\}. \quad \text{Зазначимо, що інтерпретація події } A_j \text{ має}$$

конкретний характер. Тобто, ймовірність настання події A_j (до появи B , яка може вплинути на A) визначається априорно і називається *апріорною ймовірністю* події A_j . Аналогічно, $p\{A_j | B\}$ в рівнянні (1) є *апостеріорною ймовірністю* події A_j , обчисленаю за умови, що мала місце подія B . Таким чином, можливою інтерпретацією теореми (1) є така: вона представляє собою *механізм для оновлення ступеня упевненості в тому, що відбудеться A_j в світлі нової інформації стосовно B* , яка має відношення до A_j . Первинна (апріорна) інформація $p\{A_j\}$ замінюється уточненою $p\{A_j | B\}$. З цієї точки зору теорему можна розглядати як правило навчання на основі минулого досвіду. Оновлення інформації відбувається від початкового ступеня віри (або ймовірності події) $p\{A_j\}$ до оновленої ймовірності $p\{A_j | B\}$.

Введемо декілька основних поняття, якими оперує байєсівська теорія. *Априорні ймовірності* $p\{A_j\}$ – це ступені віри (довіри), які доступні аналітику до появи та аналізу даних, які мають відношення до проблеми. У тих випадках, коли подальші дані щодо проблеми отримати неможливо, то апріорними ймовірностями можна скористатись для формульовання висновку (прийняття рішення). Тобто ТБ у таких випадках не використовується. Все, що є у розпорядженні аналітика в даному випадку – це апріорна інформація.

Апостеріорні ймовірності – це ймовірності, які ми отримуємо в результаті застосування теореми Байєса. Для того щоб розв'язок задачі був коректним, сума апостеріорних ймовірностей взаємовиключчих подій, які повністю описують деякий процес, повинна дорівнювати 1. Якщо після виконання аналізу даних за допомогою ТБ з'явились нові дані, то апостеріорні ймовірності, отримані на основі попередніх даних, необхідно використати як апріорні ймовірності для нових даних. Апостеріорній

ймовірності відповідає апостеріорна щільність, яка містить у собі апріорну та вибіркову інформацію. Вводиться також поняття *гіперпараметрів* які представляють собою параметри апріорного закону розподілу.

Визначення умовного розподілу за теоремою Байєса. Припустимо, що X – спостережувана, абсолютно неперервна змінна з функцією розподілу $f(X | \theta)$. Позначимо через Θ неспостережуваний невідомий дискретний параметр розподілу змінної X , а через $g(\theta)$ апріорну щільність розподілу Θ . В даному контексті теорема Байєса для апостеріорної щільності розподілу θ визначається виразом:

$$h(\theta | x_1, \dots, x_n) = \frac{L(x_1, \dots, x_n | \theta) g(\theta)}{\sum_{i=1}^k L(x_1, \dots, x_n | \theta_i) g(\theta_i)}, \quad (2)$$

де $L(x_1, \dots, x_n | \theta_i)$ – функція правдоподібності для θ . Форма (2) еквівалентна такій: $h(\theta | X) \propto L(x_1, \dots, x_n | \theta) g(\theta)$, де \propto – символ пропорційності. Фактичні дані можуть бути багатовимірними і можуть бути взаємно залежними, але вони повинні бути неперервними, а знаменник (2) – число, відмінне від нуля.

Приклад визначення умовного розподілу. Нехай $\vec{y}: (y_1 \dots y_n)$ вибірка з n незалежними значеннями експерименту із нормально розподіленої генеральної сукупності. Вважаємо невідомим математичне сподіванням μ , але відома дисперсію σ_0^2 . Необхідно отримати апостеріорну щільність для оцінки μ : $f(\mu / \vec{y}, \sigma_0^2) \sim f(\vec{y} / \mu, \sigma_0^2) * f(\mu)$. Оскільки кожна величина $y(k)$ розподілена нормально, то функція правдоподібності задається так:

$$\begin{aligned}
f(\vec{y}/\mu, \sigma_0^2) &= \prod_{i=1}^n f(y_i/\mu, \sigma_0^2) = (2\pi\sigma_0^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (y_i - \mu)^2\right] \\
&= (2\pi\sigma_0^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_0^2} [vs^2 + n(\mu - \tilde{\mu})^2]\right]
\end{aligned}, \quad (3)$$

де $v = n - 1$; $\tilde{\mu} = \frac{1}{n} * \sum_{i=1}^n y_i$ — вибіркове середнє; $s^2 = \frac{1}{v} \sum_{i=1}^n (y_i - \tilde{\mu})^2$ —

вибіркова дисперсія. Вираз під експонентою у другому рядку формули (3) отримано так:

$$\sum_{i=1}^n (y_i - \mu)^2 = \sum_{i=1}^n ((y_i - \tilde{\mu}) - (\mu - \tilde{\mu}))^2 = \sum_{i=1}^n (y_i - \tilde{\mu})^2 + (\mu - \tilde{\mu})^2.$$

При цьому середній член $\sum_{i=1}^n (y_i - \tilde{\mu}) * (\mu - \tilde{\mu})$ після розкриття зникає, оскільки

$\sum_{i=1}^n (y_i - \tilde{\mu}) = 0$. Припустимо що апріорну інформацію про параметр можна

представити у вигляді:

$$f(\mu) = (2\pi\sigma_a^2)^{-1/2} \exp\left[-\frac{1}{2\sigma_a^2} [(\mu - \mu_a)^2]\right],$$

де μ_a та σ_a^2 — апріорні математичні сподівання та дисперсія (гіперпараметри), тобто параметри, значення яких встановлюються дослідником на основі його первинних знань. Зазначимо, що вони представляють собою реалізацією випадкових величин.

Тепер, використовуючи теорему Байєса, запишемо:

$$f(\mu/\vec{y}, \sigma_0^2) \sim f(\vec{y}/\mu, \sigma_0^2) * f(\mu) \sim \exp\left\{-\left(\frac{\sigma_a^2 + \sigma_0^2/n}{2\sigma_a^2\sigma_0^2/n}\right) * \left(\mu - \frac{\tilde{\mu}\sigma_a^2 + \mu_a \frac{\sigma_0^2}{n}}{\sigma_a^2 + \sigma_0^2/n}\right)\right\}.$$

Звідси випливає, що апостеріорна щільність так само як і апріорна,

відповідає нормальному розподілу з математичним очікуванням:

$$M\mu = \frac{\tilde{\mu}\sigma_a^2 + \mu_a \frac{\sigma_0^2}{n}}{\sigma_a^2 + \sigma_0^2/n} = \frac{\tilde{\mu} \left(\frac{\sigma_0^2}{n} \right)^{-1} + \mu_a \left(\sigma_a^2 \right)^{-1}}{\left(\frac{\sigma_0^2}{n} \right)^{-1} + \left(\sigma_a^2 \right)^{-1}}, \quad (4a)$$

та дисперсією

$$D\mu = \left(\frac{\sigma_a^2 + \sigma_0^2/n}{2\sigma_a^2\sigma_0^2/n} \right)^{-1} = \frac{1}{\left(\frac{\sigma_0^2}{n} \right)^{-1} + \left(\sigma_a^2 \right)^{-1}}. \quad (4b)$$

Значення аналогічних параметрів для багато вимірного випадку:

$$\Sigma_*^{-1} = \Sigma_0^{-1} + n\Sigma^{-1} \quad \text{та} \quad \mu_* = \Sigma_* (\Sigma_0^{-1} \mu_0 + n\Sigma^{-1} \tilde{\mu}), \quad (5)$$

де Σ_0 і μ_0 — апріорні кореляційна матриця та вектор математичного сподівання. Отримані в цьому прикладі формули використаємо при знаходженні апостеріорних розподілів. В [3] представлені аналогічні формули для більшості інших типових апріорних законів розподілу.

Покажемо властивість оцінки параметра в формулі (4a), вона представляє собою конкретне значення параметра, яке можна рекомендувати як розв'язок задачі оцінювання. В ній наявні випадкові значення оскільки $\tilde{\mu}$ залежить від вибірки, а μ_a залежить від вибору експерта. Якщо оцінка експерта незміщена, то обчислимо математичне сподівання оцінки:

$$\begin{aligned}
M \left(\frac{\tilde{\mu} \left(\frac{\sigma_0^2}{n} \right)^{-1} + \mu_a \left(\sigma_a^2 \right)^{-1}}{\left(\frac{\sigma_0^2}{n} \right)^{-1} + \left(\sigma_a^2 \right)^{-1}} \right) &= \frac{M \left(\frac{1}{n} * \sum_{i=1}^n y_i \right) \left(\frac{\sigma_0^2}{n} \right)^{-1} + M(\mu_a) \left(\sigma_a^2 \right)^{-1}}{\left(\frac{\sigma_0^2}{n} \right)^{-1} + \left(\sigma_a^2 \right)^{-1}} = \\
&= \frac{\mu \left(\left(\frac{\sigma_0^2}{n} \right)^{-1} + \left(\sigma_a^2 \right)^{-1} \right)}{\left(\frac{\sigma_0^2}{n} \right)^{-1} + \left(\sigma_a^2 \right)^{-1}} = \mu
\end{aligned}$$

Отже, оцінка буде незміщена, якщо апріорна оцінка була незміщена. Аналогічно (тільки по координатно) виконується доведення для формул з багатовимірного випадку (5).

Розглянемо числовий приклад. Нехас існує вибірка: $Y = [0,438; 1,76; -0,647; 0,551; 1,546; -1,822; 1,308; -1,049; 0,459; -1,052]$ та апрайорні значення гіперпараметрів $\sigma_a = 2,0$; $\mu_a = 0,5$. Обчислимо значення параметрів апостеріорного розподілу за формулами (4а) та (4б) і отримаємо значення параметрів закону $M\mu = 0,1578$; $D\mu = 0,0976$.

Побудуємо функцію розподілу апрайорної (пунктирна лінія) та апостеріорної щільності (неперервна лінія):

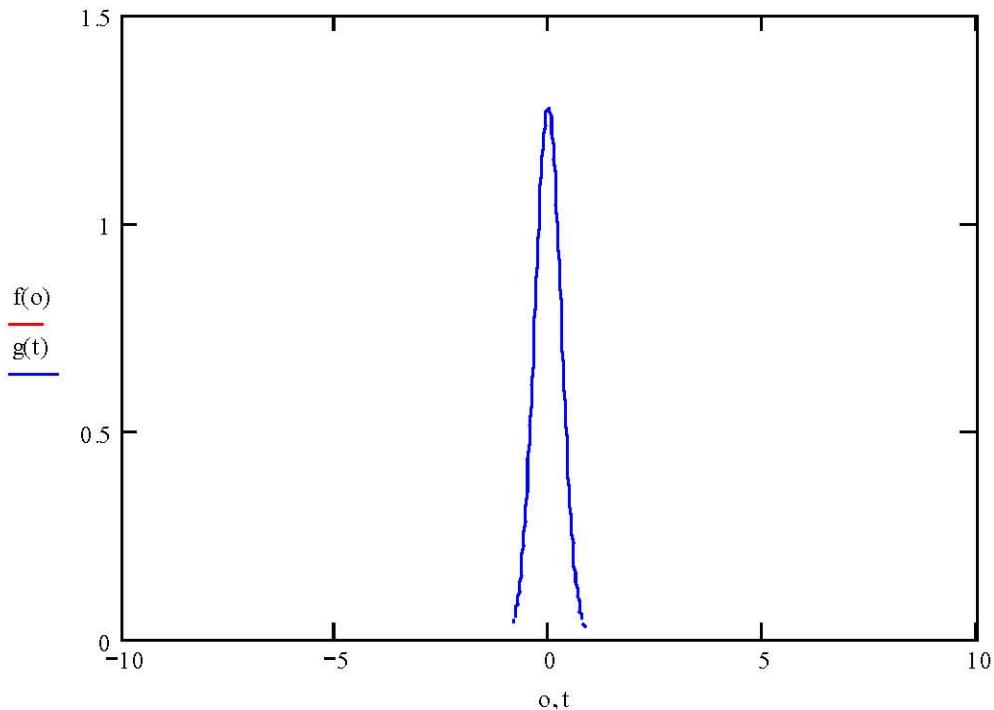


Рис. 1. Ап'єрорна та апостеріорна щільності розподілу оцінок

Отже бачимо, що суттєво зменшилась дисперсія у законі розподілу і уточнено математичне сподівання за рахунок того, що ми врахували вибіркові значення. Це означає, що при генеруванні значень ми будемо потрапляти близче до істинного значення параметра.

Застосування байєсівського підходу в методі Монте Карло. Тепер покажемо як застосовувати цей підхід для оцінювання параметрів в різних моделях. В залежності від типу моделі, на практиці реалізуються різні обчислювальні алгоритми, але за своєю суттю та підходом їх об'єднують в одну групу методів, яку називають методами Монте Карло для марковських ланцюгів (МКМЛ). Методи оцінювання цього класу основані на побудові марковського ланцюга (станами якого є вектор значень параметрів θ) із застосуванням стохастичного методу Монте-Карло.

Генерований вектор θ , є реалізацією випадкового вектора, функція щільності якого $f(\theta/y)$ корегується (уточнюються параметри) на кожному кроці, використовуючи байєсівський підхід. Так, на кожному кроці

уточнюється математичне сподівання і зменшується дисперсія розподілу $f(\theta / y)$, з якого генерується значення параметра. На кожному кроці використовується (2), а апріорним вважається значення, згенероване на попередньому кроці. Тому в результаті отримана послідовність векторів має властивість марковості.

При оцінюванні параметрів у конкретних задачах байесівська теорія застосовується разом з іншими методами оцінювання апріорних значень параметрів та методами, які дають змогу генерувати вектор Θ . Якщо модель має декілька параметрів, то логічно, що вони будуть пов'язані між собою функціональною залежністю, тобто поза діагоналлю кореляційної матриці будуть значення, відмінні від нуля і тому, в принципі, потрібно розглядати та генерувати весь вектор. У багатьох випадках це робити досить складно тому використовують спеціальні ітераційні методи які генерують вектори по-координатно, але при досить великих кількостях ітерацій розподіл генерованого вектора відповідає необхідному закону розподілу.

Генерування випадкових дискретних значень за Гіббсом. Цей метод призначений для генерування випадкового вектора $\Theta = [\theta_1 \dots \theta_p]$, розподіленого за деяким законом. Він є одним із простих та популярних МКМЛ. Метод запропонований та обґрунтований в [4, 5].

Згідно з цим методом, p -вимірний вектор розділяється на $k \leq p$ блоків, які генеруються окремо за умовним законом розподілу вигляду $\{f(\theta_j / Y, \theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_p)\}$ і таких ітерацій робиться $n \geq 1$ кроків. Тоді генеровані значення вектора Θ будуть наближатися за своїм розподілом до шуканого $f(\theta_1, \dots, \theta_p / Y)$.

Наприклад, значення розподілу вектора (X, Y) можна легко зmodелювати за допомогою алгоритму Гіббса, який використовує їх умовний розподіл. Починаючи з випадкового вибору X та Y , X моделюється з

умовного розподілу X за умови Y (тобто Y береться з попереднього кроку), а Y з умовного розподілу Y за умови наявності X . Чергуючи два умовних розподіли на послідовних кроках, генеруємо вибірку їх спільного розподілу. По мірі збільшення кількості ітерацій значення пари (X, Y) наближається до згенерованих за спільним розподілом. Необхідно згенерувати $N_0 + N$ векторів, а потім перші N_0 відкинути; тоді N реалізацій будуть розподілені за бажаним розподілом. Відзначмо, що при цьому ми можемо навіть не знати точний вигляд розподілу вектора, головне мати можливість генерувати значення з умовного розподілу. Справа в тому, що якби ми генерували одразу повний вектор, то отримали б вектор значень, координати якого пов'язані між собою деяким чином. Ці самі зв'язки і враховує повний умовний розподіл.

Приклад застосування генератора Гіббса. Розглядаємо модель лінійної регресії $y(k) = \beta_0 + \beta_1 x(k) + z(k)$, де $z(k) = \phi z(k-1) + a(k)$; $\{a(k)\} \in N(0, \sigma^2)$, $k = 1, \dots, n$. Дано вибірка з генеральної сукупності у та детерміновані значення $x(k)$.

Позначимо вектор параметрів через $\theta = [\beta_0 \ \beta_1 \ \phi \ \sigma]^T$; $\beta = [\beta_0 \ \beta_1]^T$ та $\bar{x} = [1 \ x]^T$. Будемо оцінювати вектор параметрів через генерування умовних розподілів:

$$f(\beta | \mathbf{Y}, \mathbf{X}, \phi, \sigma^2); \quad f(\phi | \mathbf{Y}, \mathbf{X}, \beta, \sigma^2); \quad f(\sigma^2 | \mathbf{Y}, \mathbf{X}, \beta, \phi).$$

Алгоритм генерування псевдовипадкової послідовності за Гіббсом

1. Робимо експертне припущення стосовно апріорних законів розподілу та параметрів законів на першому кроці для кожного параметра

$$\beta \sim N(\beta_o, \Sigma_o), \quad \phi \sim N(\phi_o, \sigma_o^2), \quad \frac{v\lambda}{\sigma^2} \sim \chi_v^2,$$

і вважаємо, що всі гіперпараметри відомі; тут χ^2_v — розподіл хі-квадрат з v ступенями свободи.

2. Оцінюємо початкові значення самих параметрів (наприклад, за МНК для $(\beta_0 \beta_1)$). Значення ϕ вважаємо відомим. На наступній ітерації його значення будемо брати з попереднього кроку. Введемо позначення

$$y_0(k) = y(k)\beta^T x_0(k) + a(k); \quad x_0(k) = x(k) - \phi x(k-1).$$

Використовуючи рівняння моделі, отримаємо:

$$y_0(k) = \beta^T x_0(k) + a(k), \quad k = 2, \dots, n.$$

Тепер, відкидаємо a_t , але врахувавши значення дисперсії його оцінки, оцінимо значення β за МНК:

$$\hat{\beta} = \left(\sum_{k=2}^n x_0(k) x_0^T(k) \right)^{-1} \left(\sum_{k=2}^n x_0(k) y_0(k) \right),$$

Ця оцінка має нормальний закон розподілу:

$$\hat{\beta} \sim N \left[\beta, \sigma^2 \left(\sum_{k=2}^n x_0(k) x_0^T(k) \right)^{-1} \right].$$

3. Обчислюємо параметри апостеріорного закону розподілу для параметрів $(\beta_0 \beta_1)$, беручи оцінки інших параметрів. Генеруємо значення з новим законом розподілу. Оскільки оцінка МНК незміщена, то математичне сподівання оцінки є шуканим значенням параметра. Тепер, використовуючи (5), отримуємо закон, з якого можна генерувати апостеріорні значення параметра:

$$(\beta | Y, X, \phi, \sigma) \sim N(\beta_*, \Sigma_*),$$

$$\Sigma_*^{-1} = \frac{\sum_{t=2}^n x_{o,t} x_{o,t}'}{\sigma^2} + \Sigma_o^{-1}, \quad \beta_* = \Sigma_* \left(\frac{\sum_{t=2}^n x_{o,t} x_{o,t}' \hat{\beta}}{\sigma^2} + \Sigma_o^{-1} \beta_o \right).$$

4. Обчислюємо параметри апостеріорного закону розподілу для параметра ϕ , беручи значення $(\beta_0 \beta_1)$ з попереднього кроку та оцінене σ . Обчислимо $z(k) = y(k) - \beta^T x(k)$ для всіх t . Тоді з виразу $z(k) = \phi z(k-1) + a(k)$,

$$z_t = \phi z_{t-1} + a_t, \quad t = 2, \dots, n.$$

аналогічно до другого кроку за МНК обчислимо оцінку:

$$\hat{\phi} = \left(\sum_{t=2}^n z_{t-1}^2 \right)^{-1} \left(\sum_{t=2}^n z_{t-1} z_t \right),$$

Використовуючи (4а) та (4б), генеруємо значення з апостеріорного нормального закону розподілу з параметрами

$$\sigma_*^{-2} = \frac{\sum_{t=2}^n z_{t-1}^2}{\sigma^2} + \sigma_o^{-2}, \quad \phi_* = \sigma_*^2 \left(\frac{\sum_{t=2}^n z_{t-1}^2 \hat{\phi}}{\sigma^2} + \sigma_o^{-2} \phi_o \right).$$

5. Обчислюємо параметри апостеріорного закону розподілу для параметра σ , беручи відомі значення інших параметрів з попереднього кроку.

Обчислюємо $a_t = z_t - \phi z_{t-1}$, $z_t = y_t - \beta^T x_t$, $t = 2, \dots, n$

та отримуємо значення σ із законом розподілу:

$$\frac{v\lambda + \sum_{t=2}^n a_t^2}{\sigma^2} \sim \chi_{v+(n-1)}^2.$$

6. Генеровані значення вважаємо оцінками параметрів та повторюємо кроки 3-6.

Кількість ітерацій може бути оцінена деякими методами [7] або обирається експертним шляхом. Загалом вона має бути досить великою $500 \square 2000$.

Інші методи генерування. Інколи для генерації випадкових одновимірних величин можна використовувати наступну теорему, але сектор

практичного застосування не широкий через нетривіальність задачі знаходження оберненої функції розподілу.

Теорема про зв'язок рівномірного розподілу з іншими.

Нехай $y = F_\xi(x)$ – функція розподілу випадкової величини ξ . Нехай існує обернена функція $x = F^{-1}(y)$ і нехай α – рівномірно розподілена випадкова величина з параметрами $[0; 1]$. Тоді η – розв'язок рівняння

$$\alpha = F_\xi(\eta)$$

– випадкова величина, розподілена за законом $F_\xi(x)$.

Доведення. Функція розподілу випадкової величини η

$$F_\eta(x) = P\{\eta < x\} = P\{F_\xi(\eta) < x\};$$

$$\eta = F_\xi^{-1}(\alpha),$$

$$\begin{aligned} F_\alpha(x) &= P\{\alpha < x\} = P\{F_\xi(\eta) < x\} = P\{\eta < F_\xi^{-1}(x)\} = \\ &= F_\eta(F_\xi^{-1}(x)) = x \quad \text{при } 0 \leq x \leq 1. \text{ При } x < 0 \text{ та } x > 1 \quad F_\xi^{-1} \text{ є невизначеною.} \end{aligned}$$

Якщо $F_1(F_2^{-1}(x)) \equiv x$, то $F_1 = F_2 \Rightarrow F_\eta(x) = F_\xi(x)$.

Алгоритм Метрополіса-Хастінгса. Розглянемо алгоритм для генерування випадкових величин із заданим розподілом [8, 9].

Припустимо, що необхідно генерувати значення деякого закону розподілу $f(\theta | X)$, але з причини наявності деяких невизначених параметрів або через складність процедури пряме генерування неможливе. Візьмемо деяку зручну для генерування функцію щільності $J_t(\theta_i | \theta_j)$, яка апроксимує необхідну щільність.

Алгоритм генерування:

1. Візьмемо випадкове значення θ_0 таке, щоб $f(\theta_0 | X) > 0$.
2. Для $k = 1, 2, \dots$

- Генеруємо кандидата вибірки з розподілу $J_k(\theta_k | \theta_{k-1})$ на ітерації k , використовуючи значення на попередньому кроці; $J_k(\theta_k | \theta_{k-1})$ – називають *розподілом стрибка*. При цьому має виконуватись умова симетричності $J_k(\theta_i | \theta_j) = J_k(\theta_j | \theta_i)$.

- Обчислюємо коефіцієнт відношення (для варіанту Метрополіса):

$$r = \frac{f(\theta_* / X)}{f(\theta_{k-1} / X)}.$$

- Покладаємо

$$\theta_k = \begin{cases} \theta_* & \text{з ймовірністю } \min(r, 1) \\ \theta_{k-1} & \text{в іншому випадку} \end{cases},$$

тобто, якщо $r > 1$, то точно приймаємо нове значення. Бачимо, що коефіцієнт вказує на те, чи наближається нове значення до математичного сподівання бажаного розподілу. Послідовність $\{\theta_k\}$ за своїм розподілом збігається до $f(\theta | X)$. Методика вибору $J_k(\theta_i | \theta_j)$ наведена у [2].

Хастінгс (1970) розвинув цей алгоритм, представивши його в такому варіанті [10]. По-перше, розподіл стрибка не обов'язково повинен бути симетричним. По-друге, умова переходу змінена на таку:

$$r = \frac{f(\theta_* / X) / J_t(\theta_* / \theta_{t-1})}{f(\theta_{t-1} / X) / J_t(\theta_{t-1} / \theta_*)},$$

а також можлива умова переходу до нового значення

$$\theta_k = \begin{cases} \theta_*, & \text{якщо } \min(r, 1) > U; \text{ де } U \text{ обране число з } (0, 1); \\ \theta_{k-1}, & \text{в іншому випадку.} \end{cases}$$

Згідно з [8, 9] оцінки, які дає цей метод, будуть незміщеними, оскільки нове значення приймається лише тоді, коли воно знаходиться в околі математичного сподівання.

Оцінювання параметрів стохастичної моделі волатильності. Тепер

ми готові до застосування викладених підходів до аналізу гетероскедастичних процесів. Повернемось до стохастичної моделі волатильності і розглянемо її у наступному вигляді

$$y(k) = \beta_0 + \beta_1 x_1(k) + \dots + \beta_p \cdot x_p(k) + \sqrt{v(k-1)} \varepsilon^s(k), \\ \log(v(k)) = \alpha_v + \beta_v \log(v(k-1)) + \sigma_v \varepsilon^v(k)$$

де $y(k)$ – часовий ряд статистичних даних, на основі яких будеться модель; $\varepsilon_s, \varepsilon_v$ – стохастичні незалежні процеси білого шуму (математичне сподівання = 0 дисперсія = 1); $\theta = \{\beta, \alpha_v, \beta_v, \sigma_v^2\}$ – вектор невідомих параметрів.

З точки зору параметрів $V(k)$ відіграє роль допоміжного параметра, але відмітимо що $V(k)$ характеризує умовну дисперсію спостерігаємої змінної:

$$Dy(k) = D(\sqrt{v}\varepsilon) = M(\sqrt{v}\varepsilon - M(\sqrt{v})M(\varepsilon))^2 = M(v)M(\varepsilon^2) = M(v)D\varepsilon = M(v).$$

Вперше МКМЛ було застосовано до цієї моделі у [2]. Далі наведемо алгоритм оцінювання параметрів, орієнтуючись на схеми в [2, 3]. За теоремою Кліфорда-Хамерслі [10] розподіл $p(\Theta, V|Y)$ повністю характеризується умовними розподілами:

$$p(\alpha_v, \beta_v | \sigma_v, v, y), p(\sigma_v^2 | \alpha_v, \beta_v, v, y) \text{ та } p(v | \alpha_v, \beta_v, \sigma_v^2, y).$$

Алгоритм оцінювання. Розділімо першу рівність моделі на $\sqrt{V(k)}$ та

пере позначимо $r_0(k) = \frac{y(k)}{\sqrt{v(k)}}; x_0(k) = \frac{1}{\sqrt{v(k)}} \begin{pmatrix} x_1(k) \\ \dots \\ x_n(k) \end{pmatrix}; \vec{\alpha} = \begin{pmatrix} \alpha_v \\ \beta_v \end{pmatrix}.$

1. В [2] пропонується брати ап'яріорний розподіл з такими параметрами $\vec{\beta} \sim N(\beta_0, A_0)$ $\vec{\alpha} \sim N(\alpha_0, C_0)$ та $\frac{m\lambda}{\sigma_v^2} \sim \chi_m^2$. Також робимо припущення щодо значення вектора \vec{V} .

2. Обчислюємо апостеріорний закон розподілу за (5) для $\vec{\beta}$. Він буде нормальним з параметрами:

$$A_*^{-1} = \sum_{t=1}^n x_{o,t} x'_{o,t} + A_o^{-1}, \quad \beta_* = A_* \left(\sum_{t=1}^n x_{o,t} r_{o,t} + A_o^{-1} \beta_o \right).$$

Щоб перейти до оцінювання наступних параметрів нам необхідно мати значення $\log(V(k))$ їх ми отримаємо з таких міркувань. Апостеріорна спільна щільність вектора:

$$p(V|\Theta, y) \propto p(y|\Theta, V) p(V|\Theta) \propto \prod_{k=1}^n p(y(k)|v(k), \Theta) p(v(k)|v(k-1), \Theta).$$

Оскільки

$$F_{\vec{\xi}}(\vec{x}) = F_{\xi_1}(x_1)^* \dots ^* F_{\xi_n}(x_n) \Rightarrow$$

$$\int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(\vec{x}) d\vec{x} = \int_{-\infty}^{x_1} f_{\xi_1}(x_1) dx_1^1 * \dots * \int_{-\infty}^{x_n} f_{\xi_n}(x_n) dx_n^n = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_{\xi_1}(x_1)^* \dots ^* f_{\xi_n}(x_n) dx_1^1 \dots dx_n^n$$

Отже маємо властивість $p(v|\Theta) \propto \prod_{t=1}^T p(v(k)|v(k-1), \Theta)$. За властивістю марковського ланцюга (V за своєю функціональною залежністю залежить лише від попереднього кроку):

$$p(v(k)|v(k-1), \Theta, y) = p(v(k)|v(k-1), v(k+1)|\Theta, y).$$

Тому далі за теоремою Байеса маємо:

$$\begin{aligned} p(v(k)|v(k-1), v(k+1), \Theta, y) &\propto p(v(k-1), v(k), v(k+1)|\Theta, y) \propto \\ &\propto p(y(k)|v(k), \Theta) p(v(k-1), v(k), v(k+1)|\Theta) \propto \\ &\propto p(y(k)|v(k), \Theta) p(v(k)|v(k-1), \Theta) p(v(k+1)|v(k), \Theta). \end{aligned}$$

Останній член $p(v(k-1)|v(k-2), \Theta)$ та члени типу $p(y(k)|v(k+1), \Theta)$ відкинуто, оскільки вони не містять $v(k)$, і являється для даного виразу константою бо попередні значення $v(k)$ вважаються такими, що прийняли деякі значення.

3. Тепер щільність волатильності, з якої необхідно генерувати

значення V , має вигляд:

$$p(v(k) | v(k-1), v(k+1), \Theta, y) \propto v^{-1/2}(k) \exp\left(-\frac{y^2(k)}{2v(k)}\right) \exp\left(-\frac{e^2(k)}{2\sigma_v^2}\right) v^{-1}(k) \times \\ \times \exp\left(-\frac{e^2(k+1)}{2\sigma_v^2}\right),$$

де $e(k) = \log(v(k)) - \alpha_v - \beta_v \log(v(k-1))$.

Генерування з такого розподілу складне, а тому використовуємо алгоритм Метрополіса-Хастінгса. В [2] запропонована щільність для генерування гамма-розподілу, оскільки перші множники якраз представляють собою функцію щільності гамма.

4. Далі, аналогічно до прикладу з регресією, оцінюємо інші параметри:

$$p(\alpha_v, \beta_v | \sigma_v, V, y) \propto \prod_{k=1}^n p(V(k) | V(k-1), \alpha_v, \beta_v, \sigma_v) p(\alpha_v, \beta_v) \propto N.$$

Параметри апостеріорного розподілу:

$$C_*^{-1} = \frac{\sum_{k=2}^n z(k) z^T(k)}{\sigma_v^2}, \quad \alpha_* = C_* \left(\frac{\sum_{k=2}^n z(k) \ln h(k)}{\sigma_v^2} + C_0^{-1} \alpha_0 \right).$$

5. За допомогою оцінених параметрів генеруємо апостеріорне значення σ з розподілу:

$$p(\sigma_v | \alpha_v, \beta_v, V, y) \propto \prod_{k=1}^n p(v(k) | v(k-1), \alpha_v, \beta_v, \sigma_v) p(\alpha_v, \beta_v) \propto IG.$$

Як визначено в [3], апостеріорний розподіл матиме вигляд:

$$\frac{m \lambda + \sum_{k=2}^n h^2(k)}{\sigma_v^2} \sim \chi_m^2; \quad h(k) = \ln v(k) - \alpha_v - \beta_v \cdot \ln v(k-1).$$

6. Кроки 2 \sqcap 5 повторюємо в циклі до тих пір, поки не досягнемо деякої кількості ітерацій необхідної для оцінок.

За таких умов алгоритм забезпечує лінійну збіжність. Тобто,

$\exists q \in (0; 1) \text{ так } k_0 > 0 \text{ таки, що } \|\theta_{k+1} - \theta\| \leq q \|\theta_k - \theta\|, \quad \forall k \geq k_0.$ Символом k позначаємо номер ітерації. Отже оцінка консистентна та із врахуванням вище наведених властивостей використаних методів незміщена.

Висновки. Розглянуто поняття нестаціонарних гетероскедастичних процесів. Наведено приклади моделей, які описують такі процеси, та вказано на основні методи оцінювання їх параметрів. З точки зору дисперсії оцінки прогнозування або оцінок параметрів логічніше віддавати перевагу умовному оцінюванню, що виконано у роботі. Для аналізу обрана стохастична модель волатильності, оскільки вона активно застосовується у багатьох напрямах в економетриці та фінансах.

Розглянута теорема Байєса та її модифікація для знаходження апостеріорного розподілу. Описано байєсівський підхід до оцінювання параметрів незалежно від вигляду моделей та типу параметрів, що оцінюються. Для реалізації піходу на практиці важливою задачею постає генерування випадкових значень апостеріорного розподілу. Розглянуто методи дискретизації Гіббса та Метрополіса-Гастінгса. Як приклад, наведено приклад застосування генератора Гіббса до моделі лінійної регресії з часовим рядом. Наведено розгорнутий алгоритм оцінювання для моделі стохастичної волатильності, який використовує різні методи генерування та способи визначення апостеріорного розподілу. Показано, що алгоритм надає незміщені та консистентні оцінки для параметрів моделі і для волатильності.

Література

1. Harvey A. C., Ruiz E., Shephard N. Multivariate Stochastic Variance Models // Review of Economic Studies, vol. 61, 1994, pp. 75-93.
2. Jacquier E., Polson N. G., Rossi P. E. Bayesian Analysis of Stochastic Volatility Models (with discussion) // Journal of Business and Economic Statistics, vol. 12. 1994, pp. 87-103.

3. Tsay L. Analysis of financial time series. – New York: Wiley series in probability and statistics, 2002. – 455 p.
4. Geman, S., Geman D. Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images // IEEE Transactions on pattern analysis and machine intelligence, 1984, 6, pp. 721-741.
5. Gelfand A. E., Smith A.F.M. Sampling Based Approaches to Calculating Marginal Densities // Journal of the American Statistical association 1990, vol. 85, pp. 398-409.
6. Roberts G. O., Smith A.F.M. Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis Hastings Algorithms // Stochastic Processes and Applications 1994, vol. 49, pp. 207-216.
7. Brooks S.P., Gareth O. Roberts Assessing Convergence of Markov Chain Monte Carlo Algorithms. – School of Mathematics, University of Bristol, 1997, 27 p.
8. N. Metropolis, S. Ulam, The Monte Carlo Method // J. Amer. statistical assoc. 1949, Vol. 44, № 247, pp. 335—341.
9. Metropolis,N., A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, "Equation of State Calculations by Fast Computing Machines", *J. Chem. Phys.*,21, 6, pp. 1087-1092, 1953.
- 10.Besag, J. (1974). ``Spatial interaction and the statistical analysis of lattice systems" // *Journal of the Royal Statistical Society, Series B*, 1974, vol. 36, pp. 192–236.