شبیه‌سازی فرآیندهای گاوسی مانا

علیرضا عرب‌پور*

چکیده

در این مقاله یک روش جالب و سریع برای شبیه‌سازی فرایندهای گاوسی مانا، وقتی که تابع کووارایانس فرایند معلوم باشد، ارائه می‌دهیم. به عبارت دیگر ماتریس $n \times n$ کوواریانس معلوم است و قصد شبیه‌سازی برداری تصادفی به طول n داریم، که دارای این ماتریس کوواریانس باشد.

۱ مقدمه

روش‌های متفاوتی برای شبیه‌سازی فرایندهای گاوسی وقتی که تابع کوواریانس فرایند مشخص باشد، وجود دارد. در این مقاله یک روش جالب برای شبیه‌سازی فرایندهای گاوسی ارائه خواهیم داد. به عبارت دیگر ماتریس کوواریانس مشخص است و قصد تولید فرایند گاوسی مانا روی بازهٔ (۱ ,۰) داریم. قبل از اینکه این روش را بیان کنیم، اشاره‌ای مختصر به نمایش فوریهٔ دنباله‌های متناهی خواهیم داشت که مورد استفاده قرار می‌گیرد.

۲ نمایش فوریه

اغلب، راحت‌تر است که یک تابع را به وسیله مجموعه‌ای از توابع مقدّماتی، که یک پایه نامیده می‌شود، نمایش دهیم. یک مجموعه بسیار مفید از این توابع مقدماتی، توابع سینوسی و کسینوسی یا نمایی مختلط است.

اگر Z_1، Z_2، ...،Z_n دنباله‌ای از n عدد دلخواه باشند، در این صورت این دنباله را می‌توان به صورت یک ترکیب خطّی از توابع مثلثاتی متعامد به شکل زیر نوشت.

$$Z_t = \sum_{k=0}^{[\frac{n}{2}]} \left[a_k \cos\left(\frac{2k\pi t}{n}\right) + b_k \sin\left(\frac{2k\pi t}{n}\right)\right],$$

$$t = 1,\ 2,\ ...,\ n.$$

این معادله را سری فوریه دنباله $\{Z_t\}$ و a_k و b_k را ضرایب فوریه گویند. این سری فوریه را، می‌توان به صورت زیر نوشت.

$$Z_t = \begin{cases} \sum_{k=-\frac{n-1}{2}}^{\frac{n-1}{2}} C_k \exp(i\omega_k t), & \text{برای } n \text{ های فرد} \\ \sum_{k=-\frac{n}{2}+1}^{\frac{n}{2}} C_k \exp(i\omega_k t), & \text{برای } n \text{ های زوج} \end{cases}$$

که $\omega_k = \frac{2k\pi}{n}$ و $k = 0,\ 1,\ ...,\ [\frac{n}{2}]$. همچنین ضرایب فوریهٔ C_k به صورت زیر محاسبه می‌شوند.

$$C_k = \frac{1}{n}\sum_{t=1}^{n} Z_t \exp(-i\omega_k t).$$

۳ تبدیل فوریهٔ سریع

برای محاسبهٔ ضرایب فوریهٔ دنبالهٔ معلوم $\{Z_t\}_{t=1}^{n}$، که تبدیلات فوریه نیز نامیده می‌شوند، معمولاً از رابطهٔ بالا

*دانشجوی دکتری گروه آمار دانشگاه شهید باهنر کرمان.

استفاده می‌شود. اغلب n را زوج و $\omega_k = \frac{2k\pi}{n}$ فرض می‌کنند که $k = -\frac{n}{2}+1, \ldots, 0, \ldots, \frac{n}{2}$.

ملاحظه می‌شود که برای هر k، محاسبهٔ C_k تقریباً n حاصل‌ضرب و حاصل‌جمع مختلط لازم دارد. بنابراین محاسبهٔ مستقیم مجموعهٔ C_k ها با استفاده از این فرمول تقریباً n^2 حاصل ضرب و حاصل جمع مختلط لازم دارد. لذا برای n های بزرگ هم از لحاظ محاسبات و هم ذخیره‌سازی اطلاعات در رایانه، دچار مشکل می‌شویم. امّا یک الگوریتم کارا که به تبدیل فوریهٔ سریع (FFT) موسوم است، برای محاسبهٔ این تبدیلات فوریه، ارائه شده است. در اصل، تبدیل فوریهٔ سریع یک الگوریتم تکراری است که از خواص توابع متعامد مثلثاتی وتوابع نمایی مختلط در کاهش محاسبهٔ تبدیل فوریهٔ n نقطه‌ای، به تبدیل ساده‌تر، استفاده می‌کند.

درحالتی که $n = 2^g$، به شرح این الگوریتم می‌پردازیم.

دنباله $\{Z_t\}_{t=1}^{n}$ را در نظر بگیرید. فرض می‌کنیم $X_t = Z_{2t}$، مقادیر $\{Z_t\}$ با اندیس زوج و $Y_t = Z_{2t+1}$ مقادیر $\{Z_t\}$ با اندیس فرد را نشان دهند. در این، صورت تبدیل فوریهٔ n نقطه‌ای این دنباله با ضرایب فوریه $C_k = \frac{1}{n}\sum_{t=1}^{n} Z_t \exp(-i\omega_k t)$ را می‌توان به صورت زیر نوشت،

$$\begin{aligned} C_k &= \frac{1}{n}\sum_{t=1}^{n} Z_t \exp(-i\omega_k t) \\ &= \frac{1}{2} f_k + \frac{1}{2}\exp\left(\frac{2ik\pi}{n}\right) g_k, \end{aligned}$$

که

$$f_k = \frac{2}{n}\sum_{t=1}^{\frac{n}{2}} X_t \exp\left(\frac{-2ik\pi t}{\frac{n}{2}}\right)$$

$$g_k = \frac{2}{n}\sum_{t=1}^{\frac{n}{2}} Y_t \exp\left(\frac{-2ik\pi t}{\frac{n}{2}}\right).$$

ملاحظه می‌شود که f_k و g_k به ترتیب تبدیلات فوریه $(\frac{n}{2})$ نقطه‌ای X_t و Y_t هستند. بنابراین، از معادلهٔ $C_k = \frac{1}{2}f_k + \frac{1}{2}\exp\left(\frac{2ik\pi}{n}\right) g_k$ نتیجه می‌گیریم که تبدیل فوریهٔ n نقطه‌ای را می‌توان بر حسب ترکیب خطی تبدیلات فوریه ساده‌تر، به طول $\frac{n}{2}$ محاسبه کرد. زیرا

$$f_k = f_{k-\frac{n}{2}} \quad \text{و} \quad g_k = g_{k-\frac{n}{2}}.$$

بدین ترتیب، کافی است که f_k و g_k را برای $k = 1, 2, \ldots, \frac{n}{2}$ محاسبه کنیم. برای اطلاعات بیشتر به بلوم فیلد[2] (۱۹۷۶، فصل ۴) رجوع شود.

[2] Bloomfield

۴ شبیه سازی فرایندهای گاوسی مانا

فرض کنید بخواهیم بردار تصادفی

$$U = \left[X\left(\frac{0}{n}\right), X\left(\frac{1}{n}\right), \ldots, X\left(\frac{n-1}{n}\right)\right]^T$$

را از یک فرایند گاوسی مانا (X) با میانگین صفر و تابع کوواریانس داده شده‌ای مانند $\gamma(.)$، تولید کنیم. در این صورت $U \sim N_n(0, G)$ که

$$G = \begin{bmatrix} \gamma(0) & \gamma(\frac{1}{n}) & \cdots & \gamma(\frac{n-1}{n}) \\ \gamma(\frac{1}{n}) & \gamma(0) & \cdots & \gamma(\frac{n-2}{n}) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(\frac{n-1}{n}) & \gamma(\frac{n-2}{n}) & \cdots & \gamma(0) \end{bmatrix}. \quad (1)$$

با توجه به براکول[3] و دیویس[4] (۱۹۸۷، فصل ۴) راجع به خواص ماتریس چرخشی، گوییم ماتریس $C = [c_{ij}]_{i,j=1}^{m}$ یک ماتریس چرخشی است، اگر تابع $f(.)$ با دورهٔ تناوب m وجود داشته به گونه‌ای که،

$$c_{ij} = f(j-i), \quad i, j = 1, 2, \ldots, m$$

یعنی

$$C = \begin{bmatrix} f(0) & f(1) & \cdots & f(m-1) \\ f(m-1) & f(0) & \cdots & f(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ f(1) & f(2) & \cdots & f(0) \end{bmatrix}.$$

در این صورت C را یک ماتریس چرخشی گویند، و تبدیل فوریه سریع، مراحل تولید بردار U به شکل زیراست:

مرحله ۱— ابتدا ماتریس G را در یک ماتریس کوواریانس چرخشی $m \times m$ مانند C، محاط می‌کنیم که $m = 2^g$ و g یک عدد صحیح است به طوری که $m \geq 2n$.

مرحله ۲— با دو بار به کارگیری روش FFT، بردار تصادفی Y که

$$Y = (Y_0, Y_1, \ldots, Y_{m-1})^T \sim N_m(0, C)$$

را تولید کرده، آن وقت با طرّاحی مناسب ماتریس C در مرحله ۱، بردار تصادفی U را از بردار Y به‌دست می‌آوریم.

[3] Brockwell
[4] Davis

	n					
	۱۰۰	۲۵۰	۳۰۰	۵۰۰	۱۰۰۰	۵۰۰۰
هر c و $\alpha \leq 1$	۸	۹	۱۰	۱۰	۱۱	۱۴
$c = 0/1$ و $1/25 \leq \alpha \leq 1/75$	۱۲	۱۳	۱۳	۱۴	۱۵	۱۸
$c = 0/1$ و $\alpha = 1/99$	۱۲	۱۳	۱۳	۱۴	۱۵	۱۷
$c = 1$ و $\alpha = 1/25$	۹	۱۰	۱۲	۱۱	۱۲	۱۵
$c = 1$ و $\alpha = 1/5$	۹	۱۱	۱۱	۱۲	۱۳	۱۶
$c = 1$ و $1/75 \leq \alpha \leq 1/99$	۱۰	۱۱	۱۲	۱۲	۱۳	۱۶
$c = 10$ و $1/25 \leq \alpha \leq 1/99$	۸	۹	۱۰	۱۰	۱۱	۱۴
$c = 10$ و $\alpha = 1/99$	۸	۱۰	۱۰	۱۱	۱۲	۱۴
$c = 100$ و $1/25 \leq \alpha \leq 1/99$	۸	۹	۱۰	۱۰	۱۱	۱۴

جدول ۱: کمترین مقدار g با تابع کوواریانس (٦)، به ازای مقادیر مختلف c، n و α، به طوری که ماتریس کوواریانس C مقادیر ویژه نامنف
داشته باشد.

ماتریس چرخشی C را به شکل زیر در نظر می‌گیریم.

$$C = \begin{bmatrix} c_0 & c_1 & \cdots & c_{m-1} \\ c_{m-1} & c_0 & \cdots & c_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{bmatrix}$$

که

$$c_j = \begin{cases} \gamma(\frac{j}{n}), & 0 \leq j \leq \frac{m}{2} \\ \gamma(\frac{m-j}{n}), & \frac{m}{2} < j \leq m-1 \end{cases} \qquad (2)$$

و γ تابع کوواریانس فرایند است. اگر $m \geq 2n$، در این صورت زیر ماتریسهای $n \times n$، از گوشه بالا به طرف پایین، روی قطر اصلی ماتریس C، برابر ماتریس G می‌باشند.

قضیه ۴. فرض کنید

$$f(\omega) = \gamma(0) + 2\sum_{k=1}^{\infty} \gamma(\frac{k}{n}) \cos(2\pi k\omega), \qquad \omega \in [0,\ 1]$$

؟ چگالی طیفی فرایند تصادفی $\{X(t) :\ t \in [-1,\ 1]\}$ باشد. فرض کنید $f(\omega)$ اکیداً روی بازهٔ $[0,\ 1]$ مثبت باشد و تابع کوواریانس فرایند X، مطلقاً جمع‌پذیر باشد؛ یعنی

$$\sum_{k=0}^{\infty} \left|\gamma(\frac{k}{n})\right| < \infty\,.$$

در این صورت ماتریس

$$C = \begin{bmatrix} c_0 & c_1 & \cdots & c_{m-1} \\ c_{m-1} & c_0 & \cdots & c_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{bmatrix}$$

که

$$c_j = \begin{cases} \gamma(\frac{j}{n}), & 0 \leq j \leq \frac{m}{2} \\ \gamma(\frac{m-j}{n}), & \frac{m}{2} < j \leq m-1. \end{cases}$$

برای هر m به اندازه کافی بزرگ، همیشه مثبت است؛ یعنی مقادی
ویژه ماتریس C نامنفی هستند.

برهان:— [۴].

این الگوریتم برای تابع کوواریانس زیر

$$\gamma(t) = \exp(-c|t|^{\alpha}),$$
$$t \in (0,\ 1),\ \ c > 0,\ \ 0 < \alpha \leq 2, \qquad (3)$$

با مقادیر مختلف α، c و n اعمال شده است که نتایج آن را د
جدول (۱) ملاحظه می‌کنید.این جدول، کمترین مقدار g را ک
مقادیر ویژه ماتریس $C_{m\times m}$ $(m = 2^g)$ نامنفی باشند، به ازای α، :
و n های مختلف نشان می‌دهد. با استفاده از این جدول با تعیین ا
بردار، U را با تابع کوواریانس (٦) می‌توان به کمک الگوریتم ذک
شده شبیه‌سازی کرد. چنانچه تابع کوواریانس به شکل دیگری
باشد، اگر در شرایط قضیه ۱ صدق کند، آنگاه می‌توان شکل تابع
کوواریانس را تغییر داد.

1 Bloomfield, P. (1976), Fourier Analysis of Time Series: An Introduction, Wiley, New York.

2 Brockvell, P. J. and Davis, R. A. (1987), Time Series: Theory and Methods. Springer, New York.

3 Chan, G. (2001), An effective method for simulating Guassian random fields. http: // www.stat.uiowa.edu.

4 Feuerverger, A., Hall, P. and Wood A. T. A. (1994), Estimation of fractal index and fractal dimension of a Gaussian process by counting the number of level crossings, J. of Time Series Analysis, V.15, No. 6.

Remarks and Comments

- There are lies, darned lies, and statistical outliers.
- Statistics means never having to say you're certain.
- Statistics is the art of never having to say you're wrong.
- Variance is what any two statisticians are at. - C.J.Bradfield
- A statistician is a person who draws a mathematically precise line from an unwarranted assumption to a foregone conclusion.
- A statistician is a person who stands in a bucket of ice water, sticks their head in an oven and says "on average, I feel fine!" - K.Dunnigan
- A statistician drowned while crossing a stream that was, on average, 6 inches deep. Most people use statistics the way a drunk uses a lamp post, more for support than enlightenment.
- Are statisticians normal?
- The weather man is never wrong. Suppose he says that there's an 80% chance of rain. If it rains, the 80% chance came up; if it doesn't, the 20% chance came up! - Saul Barron