هوالحق

Stochastic Processes

Gaussian Vectors and Gaussian processes

Note 4

مهدی جعفری سیاوشانی

دانشگاه صنعتی شریف

* مطالب این درس گفتار:

- متغیرهای تصادفی گاوسی
- بردارهای تصادفی گاوسی
- فرایندهای گاوسی
- چند مثال از فرایندهای گاوسی گسسته-زمان
- قضیه ای در مورد ایستایی فرایندهای گاوسی

- مقدمه:

فرایندهای پواسن و گاوسی از نظر سادگی، زیبایی و پرکاربرد بودن شبیه هم هستند. وقتی با مسئله جدیدی مواجه می‌شویم، معمولا ابتدا از فهمی که درباره این فرایندها داریم برای درک اولیه مسئله مورد نظر استفاده می‌کنیم. در مسائلی که به گونه‌ای اثر صف (queue) در آنها برجسته است از فهممان در مورد فرایند پواسن استفاده می‌کنیم و در مورد مسائلی که اثر نویز در آنها برجسته است فهم فرایند گاوسی به کمک ما می‌آید.

در این درس گفتار به طور مختصر به معرفی بردارهای تصادفی گاوسی و فرایندهای گاوسی می‌پردازیم. معرفی ما در اینجا خیلی سطحی است و خصوصا به بررسی خواص فرایندهای گاوسی پیوسته در زمان و نویز سفید گاوسی و ... نمی‌پردازیم. برای مباحث عمیق‌تر در این رابطه به کتاب Gallager فصل 3 مراجعه کنید.

برای شروع ابتدا خیلی مختصر به معرفی متغیر تصادفی گاوسی و برخی از خواص آن می‌پردازیم.

- متغیر تصادفی گاوسی

طبق تعریف متغیر تصادفی X توزیع گاوسی دارد اگر که تابع چگالی احتمال آن به صورت زیر باشد:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}$$

می‌توان چک کرد که تابع فوق واقعا تابع چگالی احتمال است (انتگرال آن 1 می‌شود) و امید ریاضی و واریانس این متغیر تصادفی به ترتیب μ و σ^2 است.

برای اینکه احساسی از نحوه کوچک شدن دنباله تابع چگالی احتمال گاوسی داشته باشیم، توجه می‌کنیم که احتمال‌های زیر را داریم:

$$\mathbb{P}[\,|X-\mu| > \sigma\,] = 0.318$$

$$\mathbb{P}[\,|X-\mu| > 3\sigma\,] = 0.0027$$

$$\mathbb{P}[\,|X-\mu| > 5\sigma\,] = 2.2 \times 10^{-12}$$

معمولا متغیر تصادفی با توزیع گاوسی را به صورت زیر نمایش می‌دهند:

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

تابع مولد گشتاور $\phi_X(r)$ یک متغیر گاوسی $X \sim \mathcal{N}(\mu, \sigma^2)$ برابر است با:

$$\phi_X(r) = \mathbb{E}[e^{rX}] = e^{r\mu + \frac{r^2\sigma^2}{2}}$$

یادآوری: در مورد تابع مولد گشتاور قضیه یکتایی زیر را داریم (نگاه کنید به کتاب Saeed Ghahramani به نام Fundamentals of Prob.، قضیه 10.2).

- قضیه: فرض کنید X و Y دو متغیر تصادفی با توابع مولد گشتاور $\phi_X(r)$ و $\phi_Y(r)$ باشند. اگر برای یک $\delta > 0$ داشته باشیم $\phi_X(r) = \phi_Y(r)$ برای هر $r \in (-\delta, \delta)$، آنگاه X, Y توزیع یکسانی دارند.

- بردارهای تصادفی گاوسی

(در ادامه فرض می‌کنیم اگر ذکر نشود، بردارها ستونی هستند).

- تعریف (ماتریس کواریانس): فرض کنید $\underline{X} = (X_1, \dots, X_n)^T$ یک بردار تصادفی متشکل از n متغیر تصادفی باشد. ماتریس کواریانس $\underline{\underline{C}}$ بردار تصادفی $\underline{X}$ یک ماتریس $n \times n$ است که به صورت زیر تعریف می‌شود:

$$C_{ij} = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] = C_{ji}$$

برای بردارهای تصادفی نیز می‌توانیم تابع مولد گشتاور تعریف کنیم. برای بردار تصادفی $\underline{X}$ با طول n تابع مولد گشتاور را به صورت زیر تعریف می‌کنیم:

$$\phi_X(\underline{r}) = \mathbb{E}[e^{\underline{r}^T \underline{X}}] = \mathbb{E}[e^{r_1 X_1 + \dots + r_n X_n}]$$

که $\underline{r} = (r_1, \dots, r_n)^T \in \mathbb{R}^n$. تابع مولد گشتاور n متغیره که در بالا تعریف شده است ممکن است برای همه مقادیر $\underline{r}$ وجود نداشته باشد (مشابه حالت یک بعدی).

در ادامه می‌بینیم که تابع مولد گشتاور یک بردار گاوسی در همه جا تعریف شده است.

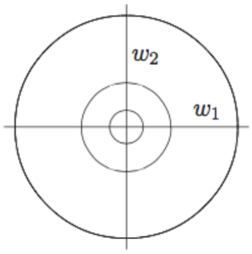
- بردارهای تصادفی گاوسی iid با توزیع استاندارد

یک بردار $\underline{W}$ از متغیرهای iid W_j را در نظر بگیرید ($1 \le j \le n$) که $W_j \sim \mathcal{N}(0,1)$.

به علت مستقل بودن W_j ها توزیع مشترک $\underline{W}$ برابر است با:

$$f_W(\underline{w}) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left(\frac{-w_1^2 - \dots - w_n^2}{2}\right) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left(\frac{-\underline{w}^T \underline{w}}{2}\right)$$

تابع چگالی مشترک $\underline{W}$ در نقطه $\underline{w}$ به مربع فاصله نقطه $\underline{w}$ یعنی $\underline{w}^T\underline{w}$ از مبدا مختصات بستگی دارد. بنابراین $f_{W}(\underline{w})$ مطابق شکل حول مبدا مختصات تقارن کروی دارد.



Equi-probability contours for an IID Gaussian 2-rv.

تابع مولد گشتاور $\underline{W}$ به سادگی به صورت زیر محاسبه می‌شود:

$$\phi_{W}(\underline{r}) = \mathbb{E}\left[e^{\underline{r}^T\underline{W}}\right] = \mathbb{E}\left[\exp(r_1 W_1 + \cdots + r_n W_n)\right]$$

$$= \mathbb{E}\left[\prod_{j=1}^{n} \exp(r_j W_j)\right] = \prod_{j=1}^{n} \mathbb{E}\left[\exp(r_j W_j)\right]$$

$$= \prod_{j=1}^{n} \exp\left(\frac{r_j^2}{2}\right) = \exp\left[\frac{\underline{r}^T\underline{r}}{2}\right]$$

که در جابجایی ترتیب ضرب و امید ریاضی از استقلال W_j ها استفاده کرده‌ایم.

- بردارهای تصادفی مشترکا گاوسی (Jointly-Gaussian random vectors)

$\{Z_1,\dots,Z_n\}$ یک مجموعه از متغیرهای مشترکا گاوسی با امید ریاضی صفر است و $\underline{Z}=(Z_1,\dots,Z_n)^T$ یک بردار گاوسی با امید ریاضی صفر است، اگر برای یک مجموعه متناهی از متغیرهای تصادفی گاوسی iid $W_i \sim N(0,1)$، $1 \leq i \leq m$، هر متغیر Z_j را بتوانیم به صورت زیر بنویسیم:

$$Z_j = \sum_{l=1}^{m} a_{jl} W_l$$

$$\underline{Z} = \underline{\underline{A}}\,\underline{W}$$

که $\underline{\underline{A}} \in \mathbb{R}^{n\times m}$. به صورت کلی‌تر، $\underline{U}=(U_1,\dots,U_n)^T$ یک بردار گاوسی است اگر که

$$\underline{U} = \underline{Z} + \underline{\mu}$$

به طوری‌که $\underline{Z}$ یک بردار گاوسی با امید ریاضی صفر است و $\underline{\mu} \in \mathbb{R}^n$.

از قبل می‌دانیم که ترکیب خطی تعداد متغیر تصادفی گاوسی مستقل از هم توزیع گاوسی دارد، بنابراین هرکدام از Z_i ها و U_i ها توزیع گاوسی دارند. نکته‌ای که در تعریف مشترکا گاوسی بودن مهم است اینست

که هر z_i ها (u_i ها) ترکیب خطی یک مجموعه مشترک از متغیرهای گاوسی مستقل هستند.

• قضیه: فرض کنید $\underline{Z} = (Z_1, \dots, Z_n)^T$ یک بردار گاوسی با امید ریاضی صفر باشد. اگر $\underline{Y} = (Y_1, \dots, Y_k)^T$ یک بردار تصادفی باشد که به صورت $\underline{Y} = \underline{\underline{B}}\,\underline{Z}$ تعریف شده باشد، در این صورت $\underline{Y}$ هم یک بردار گاوسی با امید ریاضی ۰ است.

• نتیجه: اگر $\underline{Z} = (Z_1, \dots, Z_n)^T$ یک بردار گاوسی با امید ریاضی ۰ باشد برای هر بردار حقیقی $\underline{a} \in \mathbb{R}^n$، ترکیب خطی $\underline{a}^T \underline{Z}$ یک متغیر تصادفی گاوسی با امید ریاضی ۰ است.

• مثال: فرض کنید $Z_1 \sim \mathcal{N}(0,1)$ و X مستقل از Z_1 باشد که مقادیر ± 1 را با احتمال مساوی اختیار می‌کند. در این صورت $Z_2 \sim \mathcal{N}(0,1)$ و $\mathbb{E}[Z_1 Z_2] = 0$. هرکدام از متغیرهای Z_1 و Z_2 به تنهایی گاوسی هستند ولی توزیع مشترک آنها $f_{Z_1 Z_2}(z_1, z_2)$ بر روی قطرهای $Z_2 = \pm Z_1$ پالس دارد و در بقیه جاها صفر است. بنابراین $Z_1 + Z_2$ نمی‌تواند گاوسی باشد چون مثلاً مقدار ۰ را با احتمال $\frac{1}{2}$ اختیار می‌کند.

(در نتیجه مشترکاً گاوسی با تک‌تک گاوسی بودن تفاوت دارد).

• قضیه (تابع مولد گشتاور یک بردار گاوسی با امید ریاضی صفر):

فرض کنید $\underline{Z}$ یک بردار گاوسی با امید ریاضی ۰ باشد که ماتریس کوواریانس آن $\underline{\underline{C}}$ باشد. تابع مولد گشتاور $\underline{Z}$ برابر است با:

$$\phi_Z(\underline{r}) = \mathbb{E}[\exp(\underline{r}^T \underline{Z})] = \exp\left[\frac{\underline{r}^T \underline{\underline{C}}\, \underline{r}}{2}\right]$$

اثبات: برای هر بردار $\underline{r} \in \mathbb{R}^n$ تعریف می‌کنیم: $X = \underline{r}^T \underline{Z}$. از قضایای بالا می‌دانیم که X یک متغیر گاوسی با امید ریاضی ۰ است که تابع مولد گشتاور آن برابر است با:

$$\phi_X(s) = \mathbb{E}[e^{sX}] = \exp\left(\frac{\sigma_X^2 s^2}{2}\right)$$

بنابراین برای هر $\underline{r}$ داریم:

$$\phi_Z(\underline{r}) = \mathbb{E}[\exp(\underline{r}^T \underline{Z})] = \mathbb{E}[\exp(X)] = e^{\frac{\sigma_X^2}{2}}$$

حال برای واریانس X داریم:

$$\sigma_X^2 = \mathbb{E}[|\underline{r}^T \underline{Z}|^2] = \mathbb{E}[\underline{r}^T \underline{Z}\, \underline{Z}^T \underline{r}] = \underline{r}^T \mathbb{E}[\underline{Z}\, \underline{Z}^T]\, \underline{r} = \underline{r}^T \underline{\underline{C}}\, \underline{r}$$

■

برای حالت کلی تر بردار گاوسی $\underline{U}$ با میانگین غیر صفر $\underline{\mu}$ داریم:

$$\phi_{\underline{U}}(\underline{r}) = \exp\left(\underline{r}^T \underline{\mu} + \frac{\underline{r}^T \underline{C}\, \underline{r}}{2}\right)$$

از آنجایی که تابع مولد گشتاور یک بردار تصادفی تابع توزیع مشترک متغیرهای تصادفی تشکیل دهنده بردار را به صورت یکتا مشخص می‌کند، بنابراین نتایج بالا مشخص می‌کنند که توزیع مشترک یک بردار تصادفی گاوسی به طور کامل توسط بردار امید ریاضی $\underline{\mu}$ و ماتریس کواریانس $\underline{C}$ مشخص می‌شود.

- تابع چگالی احتمال یک بردار گاوسی

فرض کنید بردار گاوسی $\underline{U}$ ماتریس کواریانس معکوس پذیر $\underline{C}$ و بردار امید ریاضی $\underline{\mu}$ را داشته باشد.

در این صورت برای تابع چگالی احتمال مشترک $\underline{U}$ داریم:

$$f_{\underline{U}}(\underline{u}) = \frac{1}{\sqrt{(2\pi)^n \det(\underline{C})}} \exp\left[-\frac{1}{2}(\underline{u}-\underline{\mu})^T \underline{C}^{-1} (\underline{u}-\underline{\mu})\right]$$

در این صورت می‌نویسیم: $\underline{U} \sim N(\underline{\mu}, \underline{C})$

* فرایندهای گاوسی

یادآوری می‌کنیم که یک فرایند تصادفی $\{X(t); t\in\mathcal{T}\}$ یک مجموعه از متغیرهای تصادفی است، یعنی به ازای هر $t\in\mathcal{T}$ که $\mathcal{T}$ مجموعه‌ای از پارامترهاست، یک متغیر تصادفی داریم. همانطور که دیده‌ایم در مبیاری از کاربردها t بیانگر زمان است، یعنی برای هر لحظه از زمان یک متغیر تصادفی داریم.

- تعریف (فرایند گاوسی): فرایند تصادفی $\{X(t); t\in\mathcal{T}\}$ فرایند تصادفی گاوسی نامیده می‌شود اگر که به ازای هر اعداد صحیح مثبت k و هر زمان‌های $t_1,\dots,t_k\in\mathcal{T}$، مجموعه متغیرهای تصادفی $X(t_1)$، ...، $X(t_k)$ یک مجموعه از متغیرهای تصادفی مشترک گاوسی باشند.

- قضیه: برای یک فرایند گاوسی $\{X(t); t\in\mathcal{T}\}$، تابع autoCovariance $C_X(t,t')$ و امید ریاضی $\mu(t)=\mathbb{E}[X(t)]$ (برای همه مقادیر $t,t'\in\mathcal{T}$) تابع چگالی مشترک بردارهای

$$(X(t_1),\dots,X(t_k))$$

را برای هر kها و هر $t_1,t_2,\dots,t_k\in\mathcal{T}$ ها مشخص می‌کنند (یعنی به طور کامل این فرایند را توصیف می‌کنند).

اثبات: طبق تعریف بردار تصادفی $(X(t_1),\dots,X(t_k))$ برای هر $k>0$ صحیح و هر $t_1,\dots,t_k$ یک بردار گاوسی است که تابع چگالی احتمال آن توسط بردار امید ریاضی و ماتریس کوواریانس این k متغیر تصادفی مشخص می‌شود. توجه می‌کنیم با داشتن امید ریاضی فرایند $\{X(t); t\in\mathcal{T}\}$ امید ریاضی بردار بالا نیز مشخص می‌شود. همچنین عنصر (i,j) ماتریس کوواریانس بردار بالا هم برابر است با:

$$C_{ij} = C_X(t_i, t_j)\ ,\quad i,j\in\{1,\dots,k\}$$

بنابراین حکم قضیه اثبات می‌شود

■

- مثال (فرایند گاوسی iid زمان-گسسته):

فرایند تصادفی $\{W(n); n\in\mathbb{Z}\}$ را در نظر بگیرید به طوری که $\dots, W(-1), W(0), W(1), \dots$ دنباله‌ای از متغیرهای تصادفی iid با توزیع گاوسی باشند، یعنی داشته باشیم $W(n)\sim\mathcal{N}(0,\sigma^2)$.

امید ریاضی این فرایند برای هر n صفر است و autoCovariance آن برابر است با:

$$C_W(k,\ell) = \sigma^2\delta_{k,\ell}$$

با توجه به تعاریفی که قبلا بیان کرده‌ایم این فرایند ایستا از نوع ضعیف است (WSS است).

برای هر k زمان n_1 و ... و n_k تابع چگالی مشترک این فرایند برابر است با:

$$f_{W(n_1),\dots,W(n_k)}(w_1,\dots,w_k) = \frac{1}{\sqrt{(2\pi\sigma^2)^k}} \exp\left[-\sum_{i=1}^{k} \frac{w_i^2}{2\sigma^2}\right]$$

این فرایند خیلی شبیه بردار گاوسی iid است، با این تفاوت که در اینجا بی‌نهایت متغیر تصادفی (بی‌نهایت بعد) داریم.

- مثال (فرایند جمع گاوسی زمان-گسسته):

فرایند تصادفی $\{S(n)\,;\, n\geq 1\}$ را که به صورت زیر از روی فرایند گاوسی iid زمان-گسسته بدست می‌آید، در نظر بگیرید:

$$S(n) = W(1) + \dots + W(n)$$

می‌توانیم بردار $(S(1),\dots,S(n))^T$ را به عنوان یک تبدیل خطی از بردار $(W(1),\dots,W(n))^T$ در نظر بگیریم. در این صورت S_1 و ... و S_n یک مجموعه از متغیرهای مشترکاً گاوسی با امید ریاضی صفر است. از آنجایی که این حرف برای هر $n\geq 1$ درست است، پس $\{S(n)\,;\, n\geq 1\}$ یک فرایند گاوسی است.

اگر $k\leq l$ برای auto Covariance این فرایند داریم:

$$C_S(k,l) = \mathbb{E}\left[\sum_{i=1}^{k} W_i \sum_{j=1}^{l} W_j\right] = \sum_{i=1}^{k} \mathbb{E}[W_i^2] = k\sigma^2$$

با استدلال مشابهی حالت $k>l$ هم قابل محاسبه است. بنابراین برای حالت کلی داریم:

$$C_S(k,l) = \min[k,l]\,\sigma^2$$

توجه کنید که این فرایند حتی از نوع ضعیف هم ایستا نیست (یعنی حتی WSS هم نیست).

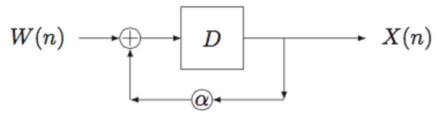
- مثال (فرایند گاوس-مارکف زمان گسسته):

فرض کنید α یک عدد حقیقی باشد به طوری که $|\alpha|<1$ و فرایند تصادفی $\{X(n)\,;\, n\in\mathbb{Z}^+\}$ را در نظر بگیرید که بر حسب فرایند گاوسی iid $\{W(n)\,;\, n\in\mathbb{Z}\}$ مثال قبل به صورت زیر تعریف شده باشد:

$$X(n+1) = \alpha X(n) + W(n), \quad n\in\mathbb{Z}^+ \,;\quad X(0)=0$$

می‌توان تصور کرد این فرایند توسط شکل بعد پیاده‌سازی شده باشد. با استفاده بازگشتی از عبارت بالا یا با استفاده از شکل داریم:

$$X(n) = W(n-1) + \alpha W(n-2) + \dots + \alpha^{n-1} W(0) \tag{3.55}$$



Schematic of the generation of $\{X(n); n \geq 1\}$ from $X(0) = 0$ and $\{W(n); n \geq 0\}$. The element D is a unit delay. It can be seen from the figure that X_{n+1} depends probabilistically on the past history $X_1, \ldots, X_n$ only through X_n. This is called a Gauss–Markov process, and the sample value x_n of X_n is called the *state* of the process at time n. This process differs from the Markov processes developed in Chapters 4, 6, and 7 in the sense that the state is an arbitrary real number rather than a discrete value.

در این مثال نیز فرایند $\{X(n) ; n \geq 1\}$ با یک تبدیل خطی از روی فرایند $\{W(n) ; n \geq 0\}$ بدست می‌آید.
چون $\{W(n) ; n \geq 0\}$ یک فرایند گاوسی است پس $\{X(n) ; n \geq 1\}$ هم یک فرایند گاوسی است (امید ریاضی آن صفر است). پس مشخصات آماری این فرایند توسط autocovariance آن مشخص می‌شود.
می‌توان نشان داد (به تمرین 3.18 کتاب Gallager نگاه کنید) که autocovariance این فرایند برابر است با:

$$C_X(n, n+k) = \mathbb{E}[X(n) X(n+k)] = \frac{\sigma^2 (1-\alpha^{2n}) \alpha^k}{1-\alpha^2}, \quad k \geq 0$$

بحث: چون $|\alpha|<1$ و ضرایب α^k در رابطه (3.55) به صورت هندسی با k کم می‌شوند پس برای n های بزرگ خیلی فرقی نمی‌کند که در محاسبه جمع فوق بر روی جمله $\alpha^{n-1} W(0)$ متوقف شویم یا جمع را با اضافه کردن جملات $\alpha^n W(-1)$ ، $\alpha^{n+1} W(-2)$ و ... ادامه دهیم.

با توجه به عبارت autocovariance هم داریم:

$$\lim_{n\to\infty} \mathbb{E}[X(n) X(n+k)] = \frac{\sigma^2 \alpha^k}{1-\alpha^2}$$

این رابطه هم نشان می‌دهد اگر به اندازه کافی از زمان شروع فرایند گذشته باشد، زمان شروع آن دیگر مهم نیست.

بنابراین عملا همان فرایند بالا را می‌توانیم بر روی همه اعداد صحیح تعریف کنیم:

$$X(n+1) = \alpha X(n) + W(n), \qquad \forall n \in \mathbb{Z}$$

یا به عبارت دیگر:

$$X(n) = \sum_{i=1}^{\infty} \alpha^{i-1} W(n-i)$$

در این صورت این فرایند دوم WSS است.

در ادامه نشان خواهیم داد که یک فرایند گاوسی WSS الزاما SSS هم است.

- قضیه: یک فرایند گاوسی $\{X(t); t\in T\}$ (که T برابر $\mathbb{Z}$، $\mathbb{R}$، $\mathbb{Z}^+$ یا $\mathbb{R}^+$ باشد، تا بتوانیم به راحتی جابجایی در زمان را تعریف کنیم) ایستا است (SSS است) اگر و فقط اگر:

$$E[X(t)] = E[X(0)], \quad \forall t\in T \qquad ①$$

$$C_X(t, t+\tau) = C_X(0,\tau) \triangleq C_X(\tau), \quad \forall t,\tau\in T \qquad ②$$

اثبات: اگر فرایند مورد نظر ایستا باشد که به وضوح شرایط بالا برقرار است.

برای اثبات در جهت عکس فرض می‌کنیم که شرایط ① و ② برقرار باشند. برای هر k و t_1 و ... و t_k و τ می‌خواهیم ثابت کنیم توزیع مشترک $X(t_1),\dots,X(t_k)$ برابر توزیع مشترک $X(t_1+\tau),\dots,X(t_k+\tau)$ است.

چون متغیرهای فوق مشترکا گاوسی هستند پس توزیع مشترک هر مجموعه توسط امید ریاضی و ماتریس کواریانس هر مجموعه مشخص می‌شود. با استفاده از شرط ① می‌دانیم که امید ریاضی هر دو بردار از متغیرهای متناظر برابر هستند.

همچنین با استفاده از شرط ② برای هر دو زمان t_i و t_j داریم:

$$C_X(t_i, t_j) = C_X(t_i+\tau, t_j+\tau)$$

پس ماتریس کواریانس این دو بردار مورد نظر هم با هم برابر هستند. در نتیجه هر دو بردار توزیع یکسانی دارند و حکم قضیه اثبات می‌شود.

■