Гауссові випадкові вектори.

1 Теоретичні відомості

Надалі позначимо $||x|| := \sqrt{\langle x, x \rangle}$, $\langle x, y \rangle = x^T y$, де $x, y \in \mathbb{R}^n$ є вектор-стовпчиками.

Трохи про математичне сподівання та коваріацію випадкового вектора.

Математичним сподіванням випадкового вектора $\vec{\xi} = (\xi_1, \ldots, \xi_n)^T$ є вектор математичних сподівань координат:

$$E[\vec{\xi}] = (E[\xi_1], \ldots, E[\xi_n])^T$$

Зокрема лінійність математичного сподівання тут узагальнюється на рівні лінійних перетворень: $E[A\vec{\xi}] = AE[\vec{\xi}]$ для довільної матриці $A \in \mathbb{R}^{d \times n}$.

Нагадаємо, що коваріацією випадкового вектора $\vec{\xi} = (\xi_1, \ldots, \xi_n)^T$ є матриця попарних коваріацій координат:

$$\Sigma := \mathrm{Cov}(\vec{\xi}) = E[(\vec{\xi} - E[\vec{\xi}])(\vec{\xi} - E[\vec{\xi}])^T] = (\mathrm{cov}(\xi_i, \xi_j))_{i,j=1}^n.$$

Очевидно, $\Sigma^T = \Sigma$ (тобто є симетричною). Зокрема Σ є невід'ємно напіввизначеною: для всіх $\vec{u} \in \mathbb{R}^n$

$$\vec{u}^T \Sigma \vec{u} = \vec{u}^T [(\vec{\xi} - E[\vec{\xi}])(\vec{\xi} - E[\vec{\xi}])^T] \vec{u} = E[\vec{u}^T (\vec{\xi} - E[\vec{\xi}])(\vec{\xi} - E[\vec{\xi}])^T \vec{u}] = E[\langle \vec{u}, (\vec{\xi} - E[\vec{\xi}]) \rangle^2] \geq 0$$

Гауссові вектори.

Попередньо ми мали справу з нормальними випадковими величинами. Це поняття можна узагальнити до випадкових векторів.

Розглянемо деякий вектор $\mu \in \mathbb{R}^n$ та *невироджену* симетричну невід'ємно напіввизначену матрицю $\Sigma \in \mathbb{R}^{n \times n}$.

Чому такий вибір параметрів?

Випадковий вектор $\vec{\xi} = (\xi_1, \ldots, \xi_n)^T$ називають гауссовим (або нормальним) вектором з вектором середніх μ та коваріаційною матрицею Σ, якщо він має щільність розподілу

$$f_{\vec{\xi}}(\vec{x}) = \frac{(\det(\Sigma^{-1}))^{1/2}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}((\vec{x} - \mu)^T \Sigma^{-1} (\vec{x} - \mu))\right) =$$
$$= \frac{1}{(2\pi)^{n/2} (\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}((\vec{x} - \mu)^T \Sigma^{-1} (\vec{x} - \mu))\right), \ \vec{x} \in \mathbb{R}^n.$$

Позначення: $\vec{\xi} \sim N(\mu, \Sigma)$.

Насправді поняття випадкового вектора на випадок виродженої коваріаційної матриці Σ, в термінах *характеристичної функції* $\varphi_{\vec{\xi}}(\vec{x}) := E[e^{i\langle \vec{x}, \vec{\xi}\rangle}] = \exp(i\langle \vec{x}, \vec{\mu}\rangle - \frac{1}{2}\vec{x}^T\Sigma\vec{x})$. Але тут не піде мова про характеристичні функції, це згодом.

Гауссів вектор $\vec{\xi}$ називають стандартним гауссовим, якщо $\mu = \vec{0}$ та $\Sigma = I_n$, де I_n – одинична матриця.

Взагалі розрізняють випадкові вектори на ізотропні та анізотропні. Випадковий вектор $\vec{\eta} \in \mathbb{R}^n$ називають ізотропним, якщо $\mathrm{Cov}(\vec{\eta}) = \sigma^2 I_n$ для деякого $\sigma > 0$. Інакше випадковий вектор називають анізотропним.

Ізотропний гауссів вектор $\vec{\xi} \sim N(\mu, \sigma^2 I_n)$ має щільність:

$$f_{\vec{\xi}}(\vec{x}) = \frac{1}{(2\pi)^{n/2}\sigma} \exp\left(-\frac{1}{2\sigma} \cdot ||\vec{x}||^2\right)$$

Ізотропні гауссові вектори хороші тим, що їх розподіл є інваріантним відносно ортогональних перетворень (зокрема поворотів на площині / у просторі). А також відмітимо те, що координати ізотропного гауссового вектора є незалежними в сукупності.

Рекламна пауза у перетворення випадкових векторів.

Для розв'язання задач, пов'язаних із гладкими перетвореннями випадкового вектора, стане у нагоді наступна теорема.

Теорема (про щільність перетвореного вектора). Нехай ξ є n-вимірний абсолютно неперервним випадковим вектором, а функція $h : \mathbb{R}^n \to \mathbb{R}^n$ кусково взаємно-однозначна та кусково неперервно диференційовна, тобто існує таке розбиття $\mathbb{R}^n = \cup_{k\geq 1} A_k$, A_k – попарно несумісні, що для довільного $k \geq 1$ $h \in C^1(A_k)$ та на A_k існує обернена до h функція $g^k = (g_1, \ldots, g_n^k) : h(A_k) \to A_k$. Тоді вектор $\eta = h(\xi)$ має щільність

$$f_\eta(\vec{y}) = \sum_{k:\vec{y}\in h(A_k)} f_\xi(g^k(\vec{y}))|J_{g^k}(\vec{y})|,$$

де $J_{g^k}(\vec{y}) = \det\left(\frac{\partial g_i^k}{\partial y_j}(\vec{y})\right)_{i,j=1}^n$ – якобіан функції g^k.

Зауваження: $|J_{g^k}(\vec{y})| = 1/|J_h(g^k(\vec{y}))|$.

Доведення наведемо далі, а зараз приклад.

Приклад. Нехай $\vec{\xi} = (\xi_1, \xi_2)^T$ – двовимірний випадковий вектор з щільністю $f(x_1, x_2)$, а $(\rho, \varphi)^T \in (0, +\infty) \times [0, 2\pi)$ – його представлення у полярних координатах. Введемо перетворення $g : \mathbb{R}^2 \to (0, +\infty) \times [0, 2\pi)$ з декартових координат у полярні. Для g відоме обернене перетворення:

$$g^{-1}(r, \theta) = (r\cos(\theta), r\sin(\theta)),\ (r, \theta) \in (0, +\infty) \times [0, 2\pi).$$

Якобіан g^{-1} дорівнює $J_{g^{-1}}(r, \theta) = r > 0$. Отже, щільність полярних координат h можна виразити через початкову щільність згідно теореми

$$h(r, \theta) = f(r\cos(\theta), r\sin(\theta))r.$$

Доведення теореми.

Візьмемо довільну підмножину $A \in \mathbb{R}^n$. Тоді

$$P(\eta \in A) = \sum_{k=1}^{+\infty} P(h(\vec{\xi}) \in A, \xi \in A_k) = \sum_{k=1}^{+\infty} P(\vec{\xi} \in g^k(A), \xi \in A_k) =$$

$$= \sum_{k=1}^{+\infty} P(\vec{\xi} \in g^k(A) \cap A_k) = \sum_{k=1}^{+\infty} \int\limits_{g^k(A)\cap A_k} f_{\vec{\xi}}(\vec{u})d\vec{u} = |\vec{x} = h(\vec{u})| =$$

$$= \sum_{k=1}^{+\infty} \int\limits_{A\cap h(A_k))} f_{\vec{\xi}}(g^k(\vec{x}))|J_{g^k}(\vec{x})|d\vec{x} = \int\limits_A \sum_{k=1}^{+\infty} 1\{\vec{x} \in h(A_k)\} f_{\vec{\xi}}(g^k(\vec{x}))|J_{g^k}(\vec{x})|d\vec{x} =$$

$$= \int\limits_A \sum_{k\geq 1:\vec{x}\in h(A_k)} f_{\vec{\xi}}(g^k(\vec{x}))|J_{g^k}(\vec{x})|d\vec{x}$$

Оскільки A – довільне, то значить підінтегральна функція в останньому виразі задає щільність розподілу перетвореного вектора $\eta = h(\xi)$.

2 Задачі

2.1 Задача 1 (позбуваємося залежності)

Нехай $\vec{\xi} = (\xi_1, \xi_2)^T \sim N((0,0)^T, A)$, де $A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $|\rho| < 1$.

1. Довести, що вектор $\vec{\eta} = (\xi_1, \xi_2 - a\xi_1)^T$ є гауссовим вектором.
2. При яких $a \in \mathbb{R}$ координати вектора з попереднього пункту є некорельованими? Незалежними?

Розв'язання.

По суті $\vec{\eta} = A\vec{\xi}$, де $A = \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix}$.

А довільне лінійне перетворення нормального вектора є нормальним вектором. Залишається визначити параметри перетвореного вектора:

$$\begin{aligned} E[\vec{\eta}] &= E[A\vec{\xi}] = AE[\vec{\xi}] = A\vec{0} = \vec{0}, \\ \mathrm{Cov}(\vec{\eta}) &= E[(\vec{\eta} - E[\vec{\eta}])(\vec{\eta} - E[\vec{\eta}])^T] = E[A\vec{\xi}(A\vec{\xi})^T] = AE[\xi\xi^T]A^T = A\mathrm{Cov}(\vec{\xi})A^T = \\ &= \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \rho - a \\ \rho - a & 1 - 2a\rho + a^2 \end{pmatrix} \end{aligned}$$

Тобто $Var[\eta_1] = 1$, $Var[\eta_2] = 1 - 2a\rho + a^2$ та $\mathrm{cov}(\eta_1, \eta_2) = \rho - a$. Для того, щоб η_1 та η_2 були некорельованими, неважко побачити що $a = \rho$. Зокрема, за цієї умови ці координати є незалежними, оскільки тоді $\Sigma := \mathrm{Cov}(\vec{\eta}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 - \rho^2 \end{pmatrix}$, і звідси $\det \Sigma = (1 - \rho^2)$ та

$$f_{\vec{\eta}}(u_1, u_2) = \frac{1}{2\pi\sqrt{1 - \rho^2}} \cdot \exp\left(-\frac{1}{2}\left(u_1^2 + u_2^2/(1 - \rho^2)\right)\right) = \varphi(u_1) \cdot \frac{\varphi(u_2/\sqrt{1 - \rho^2})}{\sqrt{1 - \rho^2}},$$

тобто сумісна щільність розпадається на добуток щільностей розподілів координат. В записі вище $\varphi(t)$ є щільністю стандартного нормального розподілу.

Коментар. В принципі умови на a досить і витягти без протягування за собою усіх параметрів перетвореного вектора. Досить розглянути коваріацію координат:

$$\mathrm{cov}(\eta_1, \eta_2) = \mathrm{cov}(\xi_1, \xi_2 - a\xi_1) = \mathrm{cov}(\xi_1, \xi_2) - a\mathrm{cov}(\xi_1, \xi_1) = \rho - aVar[\xi_1] = \rho - a.$$

2.2 Задача 2 (перетворення Бокса-Мюллера)

Нехай $\vec{\xi} = (\xi_1, \xi_2)^T \sim N((0,0)^T, I_2)$, де I_2 – одинична матриця.

1. Знайдіть розподіл полярних координат $\vec{\nu} = (\rho, \varphi)^T$ вектора $\vec{\xi}$.
2. Нехай $u_1, u_2 \sim U(0,1)$ – незалежні випадкові величини та введемо $R = \sqrt{-2\ln(u_1)}$, $A = 2\pi \cdot u_2$. Довести, що $(R \cdot \cos(A), R \cdot \sin(A))^T =^d \vec{\xi}$.

Розв'язання.

Спочатку запишемо щільність розподілу випадкового вектора $\vec{\xi}$:

$$f_{\vec{\xi}}(u,v) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(u^2+v^2)\right) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(u^2+v^2)\right) \cdot 1_{\mathbb{R}^2}((u,v)).$$

Знайдемо представлення вектора $\vec{\xi}$ у полярних координатах. Нехай $g : \mathbb{R}^2 \to (0, +\infty) \times [0, 2\pi)$ є переходом з декартової системи координат в полярну. Неважко пригадати (переконатися), що обернене перетворення має вигляд:

$$g^{-1}(r,\theta) = (r\cos(\theta), r\sin(\theta))^T,\ (r,\theta)^T \in (0,+\infty) \times [0, 2\pi).$$

Якобіан оберненого перетворення дорівнює $J_{g^{-1}}(r,\theta) = r > 0$ при всіх (r,θ). Відображення є неперервно диференційовним та взаємно однозначним, що дозволяє скористатися теоремою про щільність перетвореного вектора:

$$f_{\vec{\nu}}(r,\theta) = f_{\vec{\xi}}(r\cos(\theta), r\sin(\theta))|r| = \frac{r}{2\pi} \exp\left(-\frac{1}{2}((r\cos(\theta))^2 + (r\sin(\theta))^2)\right) 1_{(0,+\infty)\times[0,2\pi)}((r,\theta))$$

$$= 1_{[0,2\pi)}(\theta)\frac{1}{2\pi} \cdot 1_{(0,+\infty)}(r) r \exp(-r^2/2),$$

оскільки $(\sin(x))^2 + (\cos(x))^2 = 1$ та $1_{\mathbb{R}^2}(g^{-1}(r,\theta)) = 1_{g(\mathbb{R}^2)}((r,\theta)) = 1_{(0,+\infty)\times[0,2\pi)}((r,\theta))$.

Отже, полярні координати вектора $\vec{\xi}$ є незалежними, зокрема радіус ρ має розподіл Релея:

$$f_\rho(r) = 1_{(0,+\infty)}(r) r e^{-r^2/2},$$

а кут φ має рівномірний розподіл на $[0, 2\pi)$: $f_\varphi(\theta) = 1_{[0,2\pi)}(\theta)\dfrac{1}{2\pi}$.

Тепер доведемо другий пункт задачі. Неважко помітити, що $A \sim U(0, 2\pi)$, а R має розподіл Релея (переконайтеся самостійно, або подивіться матеріали занять з випадкових величин). Зокрема R та A незалежні. Сумісна щільність матиме вигляд

$$f_{(R,A)}(r,a) = f_R(r) f_A(a) = 1_{(0,2\pi)}(a)\frac{1}{2\pi} \cdot 1_{(0,+\infty)}(r) r \exp(-r^2/2)$$

Так це ж по суті майже скрізь є сумісним розподілом стандартного нормального вектора в полярних координатах, тобто $(R,A)^T =^d \vec{\nu}$. А $(R \cdot \cos(A), R \cdot \sin(A))^T = g^{-1}(R,A)$, отже має місце твердження другого пункта задачі.

Коментар. Умова незалежності координат u_1, u_2 вагома. Інакше ми би не могли запросто взяти і подати сумісний розподіл (u_1, u_2) добутком щільностей координат, у вихідному сумісному розподілі вийшла б 'нероздільна' частина.

2.3 Задача 3 (про добуток корельованих координат)

Нехай $\vec{\xi} = (\xi_1, \xi_2)^T \sim N((0,0)^T, A)$, де $A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $|\rho| < 1$.

1. Нехай $\vec{\eta} = (\eta_1, \eta_2)^T \sim N((0,0)^T, I_2)$, де I_2 – одинична матриця. Показати наступне:
$$P(\xi_1 > 0, \xi_2 > 0) = P(\eta_1, \eta_2 > -\rho/\sqrt{1-\rho^2} \cdot \eta_1) = \frac{1}{2\pi} \cdot \left(\frac{\pi}{2} + \arcsin(\rho)\right)$$
2. Скористатися попереднім пунктом задачі, щоби отримати
$$P(\xi_1 \cdot \xi_2 < 0) = \frac{\arccos(\rho)}{\pi}$$

Розв'язання.

Спочатку покажемо, що з анізотропного вектора $\vec{\xi}$ можна отримати ізотропний $\vec{\eta}$. Пригадаємо, що таку процедуру можна отримати завдяки лінійному перетворення, як показано в першій задачі: для $A = \begin{pmatrix} 1 & 0 \\ -\rho & 1 \end{pmatrix}$ маємо
$$A\vec{\xi} = (\xi_1, \xi_2 - \rho\xi_1)^T \sim N\left(\vec{0}, \begin{pmatrix} 1 & 0 \\ 0 & 1-\rho^2 \end{pmatrix}\right).$$

Але це ще не відповідає розподілу $\vec{\eta}$: треба підкоригувати дисперсію другої координати. Домножимо другий рядок матриці A на $\sqrt{Var[\xi_2 - \rho\xi_1]}$: $A_* = \begin{pmatrix} 1 & 0 \\ -\rho/\sqrt{1-\rho^2} & 1/\sqrt{1-\rho^2} \end{pmatrix}$, тоді вже $A_*\vec{\xi} = (\xi_1, (\xi_2 - \rho\xi_1)/\sqrt{1-\rho^2})^T$ матиме такий самий розподіл, що і $\vec{\eta}$. Звідси
$$P(\xi_1 > 0, \xi_2 > 0) = P(\eta_1 > 0, \sqrt{1-\rho^2}\eta_2 + \rho\eta_1 > 0) = P(\eta_1 > 0, \eta_2 > -\rho/\sqrt{1-\rho^2}\eta_1)$$

Залишається обчислити останню імовірність шляхом інтегрування. Менш болючим підходом, напевно, буде перехід до полярних координат. Спочатку зауважимо, що ми інтегруємо на множині $A = \{(x,y) \in \mathbb{R}^2 \mid x > 0, y > -\rho/\sqrt{1-\rho^2}x\}$.

$$P(\eta_1 > 0, \eta_2 > -\rho/\sqrt{1-\rho^2}\eta_1) = \int\limits_A f_{\vec{\eta}}(\vec{u})d\vec{u} = \int\limits_0^{+\infty} \int\limits_{-\rho/\sqrt{1-\rho^2}u_1}^{+\infty} \frac{1}{2\pi} \exp\left(-\frac{1}{2}(u_1^2 + u_2^2)\right) du_2 du_1 =$$
$$= |\ (u_1, u_2) := (r\cos(a), r\sin(a)), r > 0, a \in [0, 2\pi)\ | = \int\limits_0^{+\infty} \int\limits_{\arctan(-\rho/\sqrt{1-\rho^2})}^{\pi/2} \frac{1}{2\pi} \exp(-r^2/2) da dr =$$
$$= \frac{1}{2\pi}\left(\frac{\pi}{2} + \arctan(\rho/\sqrt{1-\rho^2})\right) \times \underbrace{\int\limits_{-\infty}^{+\infty} 1_{(0,+\infty)}(r) r e^{-r^2/2} dr}_{\substack{=1,\ \text{як інтеграл по щільності} \\ \text{розподілу на } \mathbb{R}}} = \frac{1}{2\pi}\left(\frac{\pi}{2} + \arctan(\rho/\sqrt{1-\rho^2})\right).$$

Нехай $|\varphi| < \pi/2$, $\varphi := \arctan(\rho/\sqrt{1-\rho^2})$. Побачимо таке:

$$\tan(\varphi) = \frac{\sin(\varphi)}{\cos(\varphi)} = \frac{\rho}{\sqrt{1-\rho^2}} \Leftrightarrow \sin(\varphi) = \rho \Rightarrow \arctan(\rho/\sqrt{1-\rho^2}) = \arcsin(\rho).$$

Це вже дозволяє отримати результат першої підзадачі:

$$P(\xi_1 > 0, \xi_2 > 0) = \frac{1}{2\pi}\left(\frac{\pi}{2} + \arcsin(\rho)\right).$$

Для розв'язання другої частини задачі, перейдемо до доповнення

$$\begin{aligned} P(\xi_1 \cdot \xi_2 < 0) &= 1 - P(\xi_1 \cdot \xi_2 \geq 0) = 1 - 2P(\xi_1 > 0, \xi_2 > 0) = \\ &= 1 - \frac{1}{\pi}\left(\frac{\pi}{2} + \arcsin(\rho)\right) = \frac{1}{\pi}\left(\frac{\pi}{2} - \arcsin(\rho)\right) = \frac{\arccos(\rho)}{\pi}, \end{aligned}$$

що завершує розв'язок задачі. Зауважимо, ми додатково скористалися наступним:

$$\begin{aligned} P(\xi_1 < 0, \xi_2 < 0) &= P(\eta_1 < 0, \eta_2 < -\rho/\sqrt{1-\rho^2}\eta_1) = P((-\eta_1) < 0, (-\eta_2) < -\rho/\sqrt{1-\rho^2}(-\eta_1)) = \\ &= P(\eta_1 > 0, \eta_2 > -\rho/\sqrt{1-\rho^2}\eta_1) = P(\xi_1 > 0, \xi_2 > 0), \end{aligned}$$

бо $-\vec{\eta} =^d \vec{\eta}$.