

به نام خدا

* روابط آنتروپی برای بردارهای تصادفی

$$\underline{x} = [x_1 \ x_2 \ \dots \ x_n]^T$$

بردار تصادفی

$$H(\underline{x}) = H(x_1, x_2, \dots, x_n) = - \sum_{\underline{x}} P(\underline{x}) \log P(\underline{x})$$

برای زنجیره ای برای آنتروپی

$$H(\underline{x}) = H(x_1, \dots, x_n) = \sum_{i=1}^n H(x_i | x_{i-1}, \dots, x_1)$$

اثبات: رابطه زنجیره‌ای برای $P(\underline{x})$ می نویسیم

$$P(\underline{x}) = P(x_1, x_2, \dots, x_n) = P(x_1) P(x_2 | x_1) \dots P(x_n | x_{n-1}, \dots, x_1)$$

$$= \prod_{i=1}^n P(x_i | x_{i-1}, \dots, x_1)$$

$$\Rightarrow H(\underline{x}) = - \sum_{\underline{x}} P(\underline{x}) \log P(\underline{x})$$

$$= - \sum_{\underline{x}} (P(\underline{x})) \sum_{i=1}^n \log P(x_i | x_{i-1}, \dots, x_1)$$

$$= \sum_{i=1}^n \left(- \sum_{\underline{x}} P(\underline{x}) \log P(x_i | x_{i-1}, \dots, x_1) \right)$$

$H(x_i | x_{i-1}, \dots, x_1)$

$$I(x_i; y) = H(x) - H(x|y)$$

- رابطه زنجیره‌ای برای اطلاعات متقابل

$$I(\underline{x}; y) = I(x_1, x_2, \dots, x_n; y)$$

$$= H(\underline{x}) - H(\underline{x}|y) = H(x_1, x_2, \dots, x_n) - H(x_1, x_2, \dots, x_n|y)$$

$$= \sum_{i=1}^n H(x_i | x_{i-1}, \dots, x_1) - \sum_{i=1}^n H(x_i | x_{i-1}, \dots, x_1, y)$$

رابطه زنجیره‌ای $H(\underline{x})$

$$= \sum_{i=1}^n \left[H(x_i | \underbrace{x_{i-1}, \dots, x_1}) - H(x_i | \underbrace{x_{i-1}, \dots, x_1, y}) \right]$$

$$\underbrace{\hspace{15em}}_{I(x_i; y | x_{i-1}, \dots, x_1)}$$

الملاحظات متناهي شرح

$$I(x; y|z) = H(x|z) - H(x|y, z)$$

$$= H(y|z) - H(y|x, z)$$

$$= \underline{E_{p(z)} I(x; y|z=z)} = E_{p(z)} D \left(p(x, y|z) \parallel \frac{p(x|z)}{p(y|z)} \right)$$

$$H(x|y=y) = - \sum p(x|y) \log p(x|y)$$

$$H(x|Y) = E_{p(y)} H(x|Y=y) = - \sum \underbrace{p(x, y)}_{\log p(x|y)}$$

$$\Rightarrow I(x; y|z) = E_{P(z)} \left(D(P(x, y|z) \parallel P(x|z) P(y|z)) \right)$$

$$= E_{P(z)} \left(\sum_{x, y} P(x, y|z) \log \frac{P(x, y|z)}{P(x|z) P(y|z)} \right)$$

$$= \sum_z P(z) \left(\sum_{x, y} P(x, y|z) \log \frac{P(x, y|z)}{P(x|z) P(y|z)} \right)$$

$$= \sum_{x, y, z} \underbrace{P(x, y|z) P(z)}_{P(x, y, z)} \log \frac{P(x, y|z)}{P(x|z) P(y|z)} \quad (1)$$

$$= E_{P(x, y, z)} \left\{ \log \frac{P(x, y|z)}{P(x|z) P(y|z)} \right\} = H(x|z) - H(x|y, z) \quad (2)$$

$$= H(y|z) - H(y|x, z)$$

آنتروپی نسبی شرطی

برای دو تابع احتمال شرطی $P(x|y)$ و $q(x|y)$ آنتروپی نسبی شرطی

حاصل KL به صورت زیر تعریف می شود

$$\begin{aligned} D(P(x|y) \| q(x|y)) &= \sum_y P(y) \sum_x P(x|y) \log \frac{P(x|y)}{q(x|y)} \\ &= \sum \underbrace{P(x|y) P(y)}_{P(x,y)} \log \frac{P(x|y)}{q(x|y)} \\ &= \sum_{x,y} P(x,y) \log \frac{P(x|y)}{q(x|y)} \\ &= E_{P(x,y)} \left\{ \log \frac{P(x|y)}{q(x|y)} \right\} \end{aligned}$$

رابطه زنجیره‌ای برای آنتروپی منسبی

$$\underline{D(p(x,y) \parallel q(x,y)) = D(p(x) \parallel q(x)) + D(p(y|x) \parallel q(y|x))}$$

تمرین: اثبات رابطه زنجیره‌ای آنتروپی منسبی

$$\sum_{x,y} p(x,y) \log \frac{p(x,y)}{q(x,y)} \rightarrow \begin{aligned} & \overbrace{p(x,y)}^{p(x)p(y|x)} \leq p(y)p(x|y) \\ & \underbrace{q(x,y)}_{q(x)q(y|x)} \leq q(y)q(x|y) \end{aligned}$$

$$D(p(x,y) \parallel q(x,y)) = D(p(y) \parallel q(y)) + D(p(x|y) \parallel q(x|y))$$

(Jensen's inequality)

جensen نامساوی

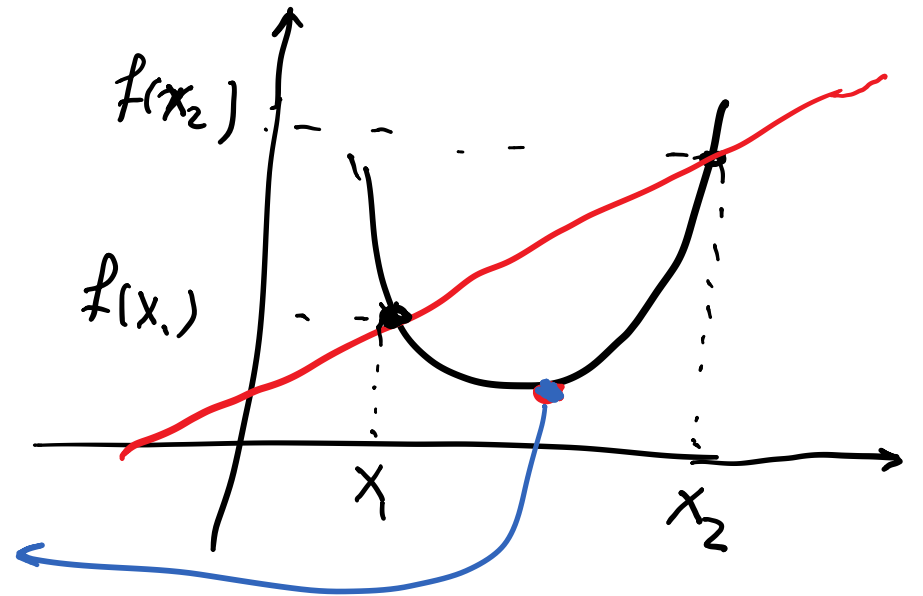
یا اگر در $f(x)$ تابع محدب باشد $[a, b]$ محدب (Convex) کریم

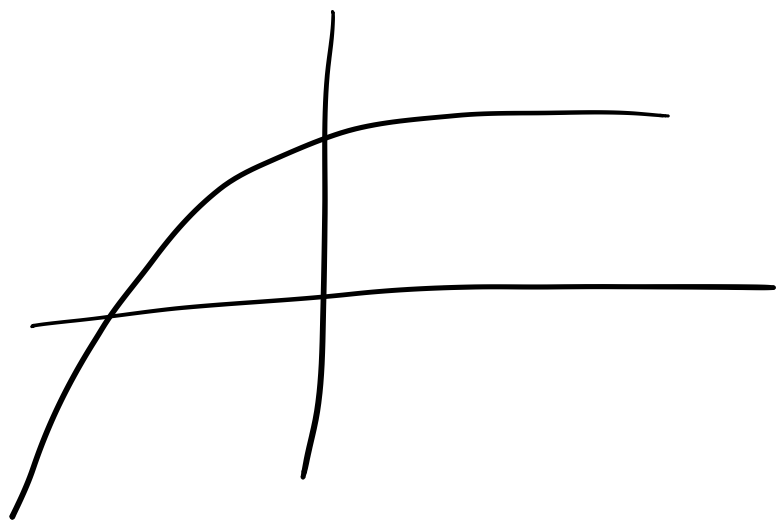
هرگاه داشته باشیم

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

$\forall x_1, x_2 \in [a, b]$, $0 \leq \lambda \leq 1$

$\lambda x_1 + (1-\lambda)x_2$

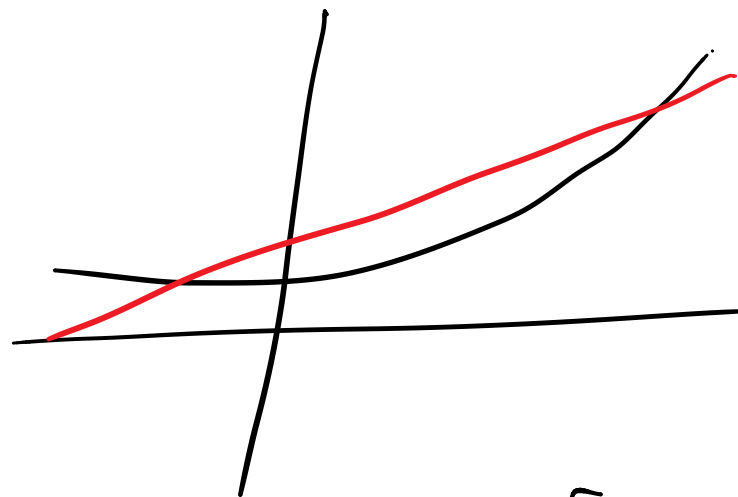




مکرب

Concave

مثال: $\log x$



متر

Convex

x^2, e^x

مثال:

تابع f را کدیه، مقعرترسیم کرده. در رابطه (*) ستادی حفظه از ای $\lambda = 0$ و $\lambda = 1$

استادیر به بی ابرقرار باشد

نقطه: تابع f مقعر یا Convex است اگر دلتا اثر مثبت دوم آن غیر منفی باشد
اثر مثبت دوم مثبت باشد تابع f آکد یا مقعر فزاینده.

* تابع f محدب یا Concave است، اگر تابع f - مقعر باشد.

نقطه: تابع f محدب یا Concave است اگر دلتا اثر مثبت دوم آن غیر مثبت باشد
اثر مثبت دوم منفی باشد، تابع f آکد یا محدب است.

قضیه: نامساوی Jensen

اگر تابع $f(x)$ مقعر باشد، x یک میانگین در نظر گرفته شود، داریم

$$f(\bar{x}) \leq \bar{f(x)}$$

تقریری سخن

$$\lambda_1 f(x_1) + \lambda_2 f(x_2) + \dots + \lambda_n f(x_n)$$

خط

اثبات به سطر

$$\bar{x} = \sum x_i \underbrace{p(x_i)}_{\lambda_i} = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$$

$$0 \leq \lambda_i \leq 1$$

$$\sum_{i=1}^n \lambda_i = 1$$

* اگر $f(x)$ محدب باشد، جهت نامساوی عکس می شود

دائر $f(x)$ اکثر استراتژیها که بایست، ستادی فتو در نتایج بدیهی رخ فزاعه دار.

نصبه: برای هر دو تابع احتمال رخزه $p(x)$, $q(x)$ هراره $D(p(x) || q(x)) \geq 0$

صفر صفت آنزوی نسبی

ستادی رخزه رخزه است آنزوی نسبی $\forall x \in X ; p(x) = q(x)$

اثبات:

$$D(p(x) || q(x)) = \sum \underline{p(x)} \log \frac{p(x)}{q(x)}$$

می فزاعیم از رابطه Jensen

که بکیریم.

$$f(t) = \log t \Rightarrow f(Ex) \geq Ef(x)$$

می رانیم که تابع t و تابعی که بکیریم.

$$-D(P(x) \parallel q(x)) = \sum_x P(x) \log \frac{q(x)}{P(x)}$$

$$= \sum_x P(x) \log q(x) - \sum_x P(x) \log P(x)$$

$$= \sum_x \underbrace{\frac{P(x)}{P(x)}}_{f(t)} \log \underbrace{\frac{q(x)}{P(x)}}_{f(t)} \leq \log \left(\underbrace{\sum_x P(x)}_1 \underbrace{\frac{q(x)}{P(x)}}_1 \right)$$

↑
jensen

$f(t) = \log t$ این Jensen برای

$$\Rightarrow -D(P(x) \parallel q(x)) \leq 0 \Rightarrow D(P(x) \parallel q(x)) \geq 0$$

$\forall x \in \mathcal{X}$

$$t = \frac{q(x)}{P(x)} = 1$$

سادس برقرار است اگر تنها اگر

مثبت $\forall x \in \mathcal{X} \quad p(x) = q(x)$

نتیجه: (قضیه) برای هر دو متغیر تصادفی x, y همواره داریم

$$I(x; y) \geq 0$$

ستادگی برقرار است که اگر $x \perp y$

اثبات: صدق

$$I(x; y) = D(P(x, y) \parallel P(x)P(y)) \geq 0$$

از قضیه مثبتی

همچنین می دانیم که ستادی برقرار است اگر و تنها اگر $\forall x, y; P(x, y) = P(x)P(y)$

یعنی $X \perp\!\!\!\perp Y$

نتیجه: $D(P(x|y) \parallel q(x|y)) \geq 0$

ستادی برقرار است. اگر و تنها اگر $P(x|y) = q(x|y)$ $\forall x, y$

اثبات: $D(P(x|y) \parallel q(x|y))$ از میانگین گیری می شود. غیر منفی حاصل

است $E_{P(y)}(D(P(x|y=y) \parallel q(x|y=y))) \geq 0$ غیر منفی خواص دارد

$$\text{نتیجه: } I(x; y | z) \geq 0$$

ساده‌ی بقرای است آوردن اگر x, y به شرط داشتن z ، مستقل باشند

استقلال شرطی x, y به شرط z

$$P(x, y | z) = P(x | z) P(y | z) \quad \text{استقلال شرطی}$$

$$I(x; y | z) = E_{P(z)} D(P(x, y | z) \parallel P(x | z) P(y | z))$$

$$= E_{P(x, y, z)} \left\{ \log \frac{P(x, y | z)}{P(x | z) P(y | z)} \right\}$$

قضیه: برای هر مجموعه‌ی X تابع احتمال $P(x)$ همواره داریم

$$H(x) \leq |x|$$

x : مجموعه تمام ستادیری است که مجموعه‌ی x ، اعضای x است. منظور از

$|x|$: تعداد اعضای مجموعه x است.

$$u(x) = \frac{1}{|x|}$$

اثبات: $u(x)$ تابع احتمال می‌باشد است داریم

$$\forall x \in X$$

داریم که انتروپی $P(x)$ ، $u(x)$ ، از نظر k لیم.

$$D(P(x) \| u(x)) = \sum_x P(x) \log \frac{P(x)}{u(x)}$$

$$= \underbrace{\sum_x P(x) \log P(x)}_{-H(x)} - \sum_x P(x) \log \underbrace{u(x)}_{\frac{1}{|X|}}$$

$$= -H(x) + \underbrace{\sum_x P(x)}_1 \log |X|$$

$$D(P(x) \| u(x)) = \log |X| - H(x) \geq 0$$

$$\Rightarrow H(x) \leq \log |X|$$

$\forall n$

ستادی برقرار است، اگر دنباله $P(n) = u(n)$

یعنی تریزج یکیزافت دارای بیشترین آنتروپی است.

قضیه: شرط ط کردن از اهام می کاهد.

$$H(x|y) \leq H(x)$$

اثبات: سی رانیم

$$I(x; y) \geq 0 \Rightarrow I(x; y) = H(x) - H(x|y) \geq 0$$

$$\Rightarrow H(x) \geq H(x|y)$$

ستادی برقرار است، اگر دنباله $x \perp y$

$$\underline{H(X|Y=y)} \not\leq H(X)$$

این رابطه در حالت کلی برقرار نیست

$$H(X|Y) = E_{P(Y)} H(X|Y=y) \geq H(X)$$

