

فاصله باتاچاریا

- اندازه گیری میزان شباهت دو توزیع احتمال
- مرتبط با ضریب باتاچاریا است که میزان همپوشانی دو نمونه نمونه یا جمعیت آماری را اندازه میگیرد.
- کاربرد: feature extraction

باتاچاریا- ماهالانوبیس

- ماهالانوبیس همون باتاچاریاست در یک حالت خاص که انحراف معیار دو کلاس یکسانه.
- مشکل: اگر میانگین ها برابر اما انحراف معیار ها نامساوی باشه، ماهالانوبیس به صفر میل میکنه.
- اما باتاچاریا بر اساس میزان تفاوت انحراف معیارها رشد میکنه.
- پس:
- باتاچاریا در کلاسبندی موفق تره.

کمی فرمول!

قاعدہ مثلثی

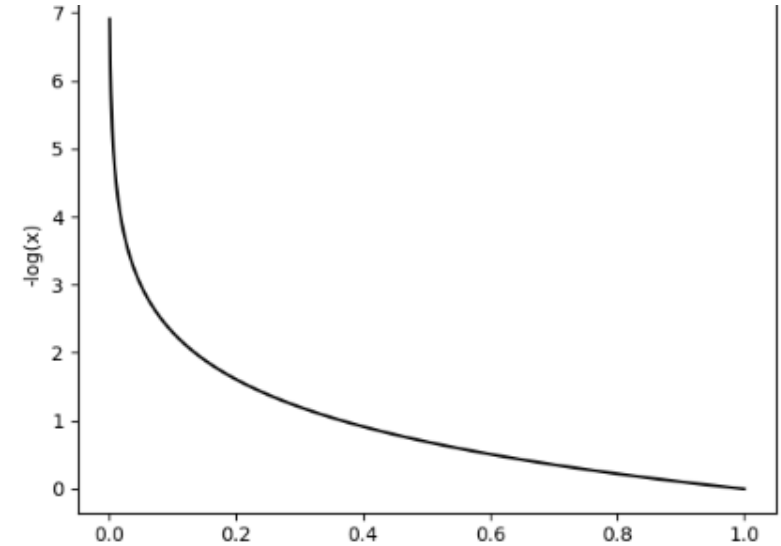
$$D_B(p, q) = -\ln(BC(p, q))$$

گسسته

$$BC(p, q) = \sum_{x \in X} \sqrt{p(x)q(x)}$$

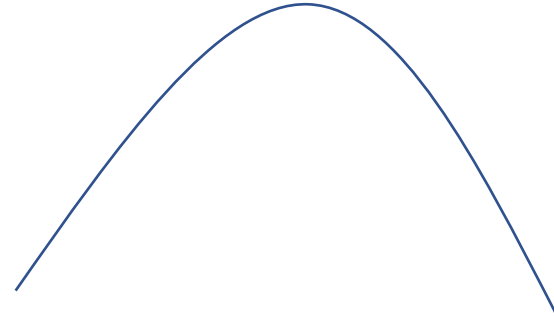
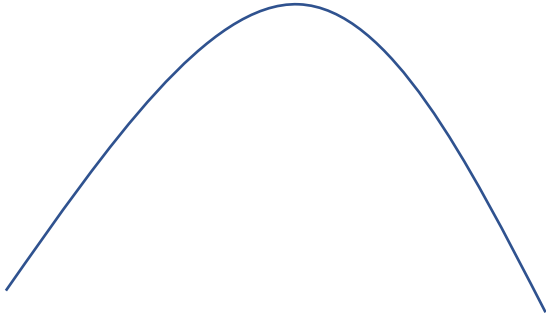
پیوسته

$$BC(p, q) = \int \sqrt{p(x)q(x)} dx$$

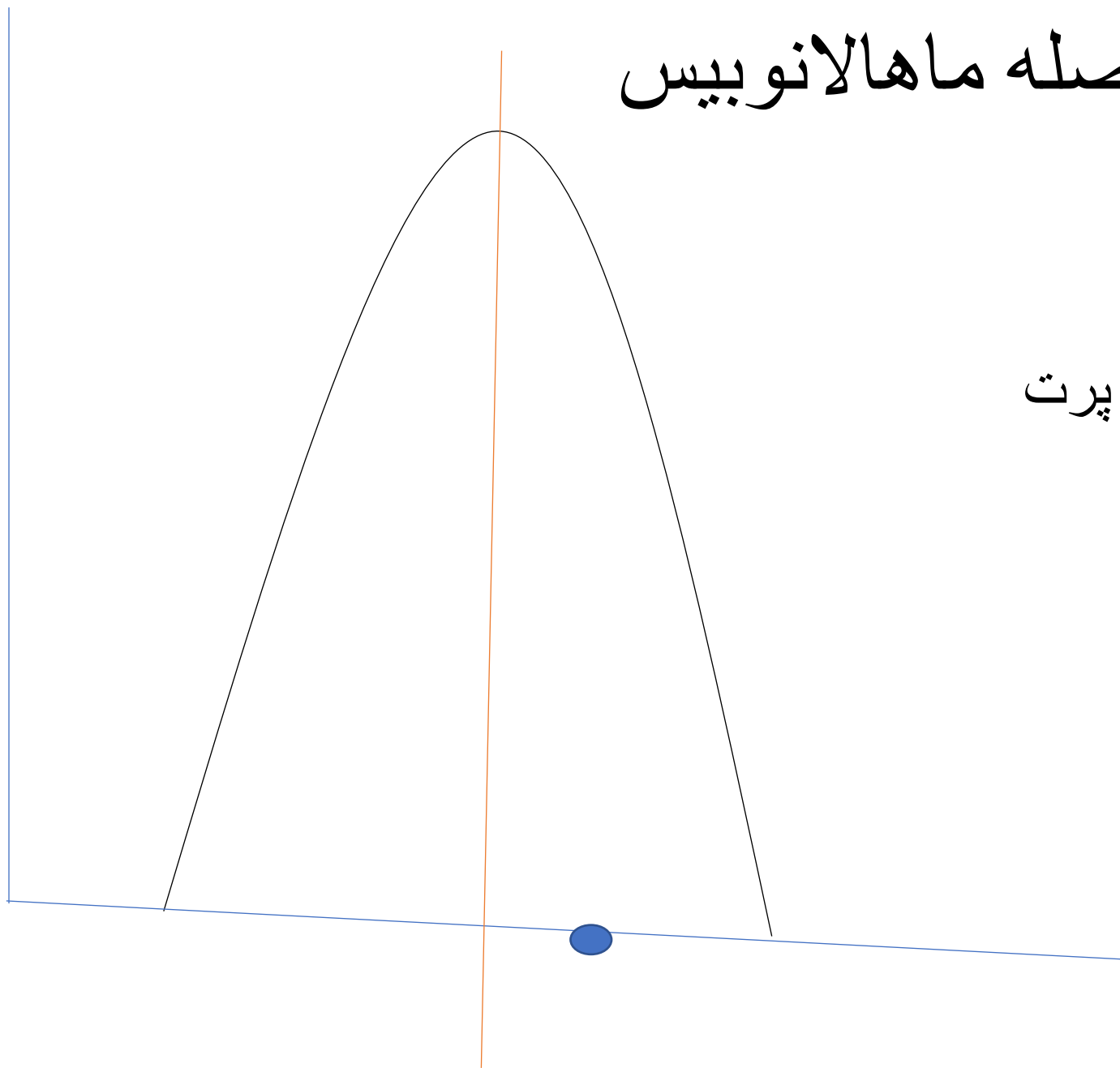


عيب پاتاچاريا

•



فاصله ماکسیمی



- اندازه بین یک نقطه با یک توزیع!
- کاربرد: خوشه بندی، تشخیص داده پرت

ماهالانوبیس همون اقلیدسیه؟

اگه محورامو (ابعادمو) اسکیل بندی کنم تا انحراف معیارها واحد) ماتریس کواریانس من همانی بشه) بشن، در فضای جدید همون اقلیدسیه.

$$d(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^N \frac{(x_i - y_i)^2}{s_i^2}},$$

کمی فرمول!

$$D_M(\vec{x}) = \sqrt{(\vec{x} - \vec{\mu})^T S^{-1} (\vec{x} - \vec{\mu})}.$$

فاصله هلینگر

- کمی سازی فاصله دو توزیع احتمال
- ظاهراً شبیه باتاچاریاست ولی خیلی نه!
- تعریف بر اساس انتگرال هلینگر
- متریک

کمی فرمول!

absolutely continuous

$$H^2(P, Q) = \frac{1}{2} \int \left(\sqrt{\frac{dP}{d\lambda}} - \sqrt{\frac{dQ}{d\lambda}} \right)^2 d\lambda.$$

Radon–Nikodym derivatives

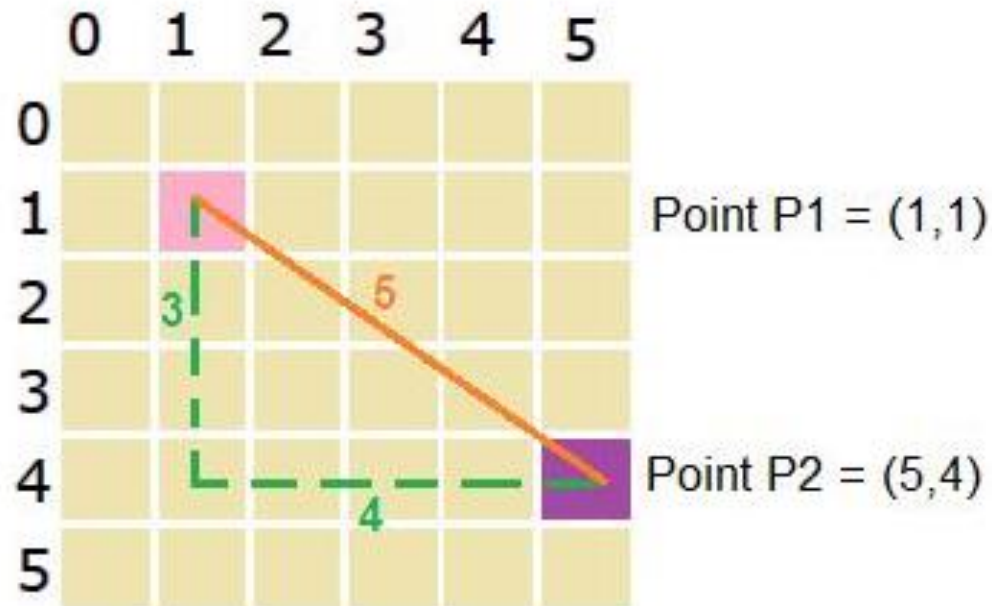
$$H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^k (\sqrt{p_i} - \sqrt{q_i})^2},$$

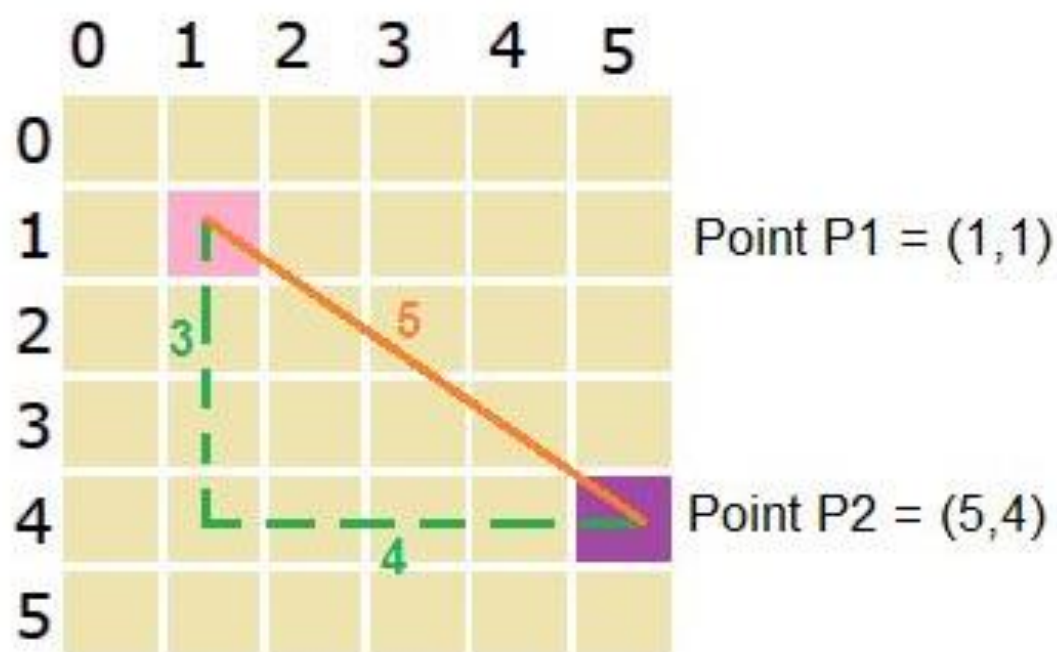
فاصله منهتن

$$d(i, j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|.$$

فرمول محاسبه فاصله منهتن Manhattan distance

- بلوک های ساختمانی
- راننده تاکسی





$$\text{Euclidean distance} = \sqrt{(5-1)^2 + (4-1)^2} = 5$$

$$\text{Manhattan distance} = |5-1| + |4-1| = 7$$

Euclidean(green) vs Manhattan(red)

مثالی از معیار فاصله منهتن

فاصله مینکوفسکی

• حالت کلی اقلیدسی و منهتن

$H \geq 1$, real

$$d(i, j) = \sqrt[H]{|x_{i1} - x_{j1}|^H + |x_{i2} - x_{j2}|^H + \dots + |x_{ip} - x_{jp}|^H},$$

فرمول معیار فاصله مینکوفسکی Minkowski distance

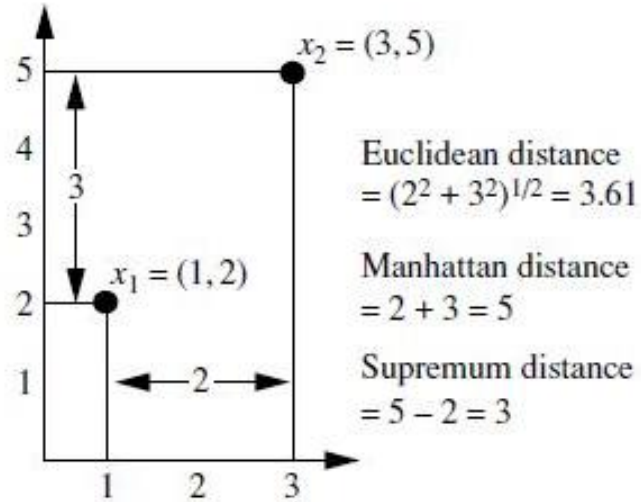
$H=1$ منهتن
 $H=2$ اقلیدسی

وزن؟؟؟

$$d(i, j) = \sqrt{w_1 |x_{i1} - x_{j1}|^2 + w_2 |x_{i2} - x_{j2}|^2 + \dots + w_m |x_{ip} - x_{jp}|^2}.$$

فرمول محاسبه فاصله اقلیدسی وزن دار

فاصله سوپریمم



مثالی از فاصله اقلیدسی و منهتن و مینکوفسکی

- حالت کلی مینکوفسکی

- فرض کن h بره به بینهایت!

- اینجوری ماکزیمم فاصله بین دو شی محاسبه میشه.

- تو بینهایت بهش uniform norm هم میگن!

$$d(i, j) = \lim_{h \rightarrow \infty} \left(\sum_{f=1}^p |x_{if} - x_{jf}|^h \right)^{\frac{1}{h}} = \max_f |x_{if} - x_{jf}|.$$

فرمول محاسبه فاصله سوپریمم

متریک چیه حالا؟

- فاصله همیشه مثبت باشه
 - فاصله هر کی از خودش صفره
 - تقارن: فاصله من با تو = فاصله تو با من
 - اصل نابرابری مثلث = جمع دو تا ضلع بیشتر از اونیکیه!
-
- این ویژگی هارو داری؟ متریکی! مثل اقلیدسی و منهتن!
 - متریک یک تابعه که اندازه ها ازش بدست میان!
 - یک سیستم استاندارد دهنده برای اندازه گیری