

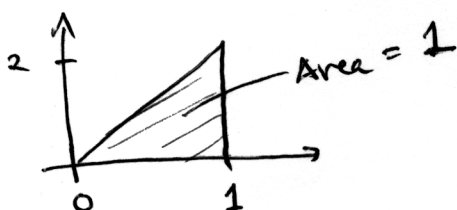
①

- a) Först klustring för att gruppera bilderna sedan lokal PCA inom varje grupp jmf. ansiktskomprimering på laboration
- b) - Att för många särdrag innebär risk för överanpassning vid träning/parameterval
- Brus i bilden reduceras normalt ~~vid~~ i lämpligt valda särdrag
- c) Ett rotations- och skalningsinvariant särdrag är relativa antalet bildpunkter ~~och~~ (area) med gråskalevärden på intervaller $[0.9, 1]$ i förhållande till antalet gråskalevärden på intervallet $[0.5, 0.9]$.

②

- a) Det går inte att på förhand säga att någon felsannolikhet e är ~~vanligare~~ vanligare än någon annan.

b)
$$\begin{matrix} N=1 \\ k=1 \end{matrix} \Rightarrow P(e|k=1, N=1) = 2e$$



- 2 c) HPD: Det kortaste intervallet som innehåller tex 95% sannolikhetsmassa

③

- a) Idén är att representera varje mönster $\bar{x}$ på formen

$$\bar{x} = t_1 \bar{b}_1 + t_2 \bar{b}_2 + \dots + t_M \bar{b}_M = \underline{\underline{B}} \underline{\underline{t}}$$

där t_i är "koordinater" i ett nytt system av "basvektorer" $\bar{b}_i$ där $M \ll N$ och N är antalet element (särdrag) i $\bar{x}$. "Basvektorena" behöve inte vara ortogonala.

Ex. PCR $\bar{b}_i$: Eigenvektor till kovariansmatrisestimat

$$\hat{C} = \frac{1}{N} \sum_{n=1}^N (\bar{x}_n - \hat{m})(\bar{x}_n - \hat{m})^T$$

som har det i :te största egenvärdet

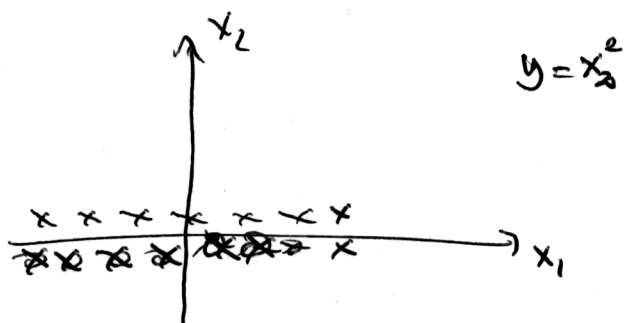
$$\Downarrow$$

$$\bar{b}_i^T \bar{b}_j = \delta_{ij} \quad (\text{ortogonala})$$

$$\Downarrow$$

$$\underline{\underline{t}}_n = \underline{\underline{B}}^T \cdot \bar{x}_n$$

b) Exempel



PCR skulle troligen ignorera x_2 och endast behålla x_1 som helt saknar koppling till y .

c) PLS: "Bas#vektorena" $\bar{b}_i$ väljs så att

1. Alla t_i maximalt korrelerade med y (responsvariabeln)
2. Alla t_i okorrelerade

Inget krav på god rekonstruktion av $\bar{x}$: söker alltså lägdimensionell representation av den del av $\bar{x}$ som korrelerar med y .

PCR: Basvektorena $\bar{b}_i$ väljs så att rekonstruktsfelet $\bar{x} - \hat{x}$ (i medel) blir så litet som möjligt oavsett hur mycket info det finns om y i $\bar{x}$.

④

a) Korsvalideringsmetoden eller m.h.a. en testmängd.

b) Nej, då behövs inte RR. Nej, CIV kunde inte ha retat delta.

c)

$$\tilde{x} = \frac{x - m}{\sigma}$$

$$E\{\tilde{x}\} = \frac{E\{x\} - m}{\sigma} = \frac{m - m}{\sigma} = 0$$

$$E\{\tilde{x}^2\} = \frac{E\{x^2 - 2xm + m^2\}}{\sigma^2}$$

$$= \frac{\sigma^2 + m^2 - 2m^2 + m^2}{\sigma^2} = 1$$

⑤

a) Mineralletning

Diagnostiserar av sjukdomar

b) Idén är att inför kostnader λ_{ij} för att välja klass j när den verkliga klassen är i . Sedan beräknas den förväntade kostnaden (i medeltal) för olika beslutsmetoder och den som ger lägst förväntad kostnad ~~blir~~ blir vald.

⑥

$$P(x|w_1) \sim N(0,1)$$

$$P(x|w_2) \sim N(1,2)$$

Optimalt: välj klass 1 om

$$P(w_1|x) > P(w_2|x)$$

$$\Downarrow$$

$$P(x|w_1)P(w_1) > P(x|w_2)P(w_2)$$

$$\text{Antag } P(w_1) = P(w_2)$$

$$\Downarrow$$

$$\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} > \frac{e^{-\frac{1}{2}\left(\frac{x-1}{2}\right)^2}}{2\sqrt{2\pi}}$$

$$-\frac{x^2 \cdot 4}{2 \cdot 4} > -\frac{1}{2} \frac{x^2 - 2x + 1}{4} = -\frac{x^2}{8} + \frac{x}{4} - \frac{1}{8}$$

Välj klass 1 om

$$\frac{3x^2}{8} - \frac{x}{4} + \frac{1}{8} > 0$$

$$7 \text{ (a)} \quad \frac{1}{\lambda_p} \underline{\underline{w_p}} = \frac{1}{\lambda_p} \frac{1}{N} \sum_{n=1}^N \bar{x}_n \bar{x}_n^T \bar{a}_p$$

$$= \sum_{n=1}^N a_{np} \bar{x}_n = \underline{\underline{X}} \bar{a}_p$$

$$a_{np} = \frac{\bar{x}_n^T \bar{w}_p}{\lambda_p \cdot N}$$

$$(b) \quad \underline{\underline{X}}^T \underline{\underline{X}} \cdot \bar{a}_p = \underline{\underline{X}}^T \cdot \bar{w}_p = \underbrace{(\underline{\underline{X}}^T \underline{\underline{X}})^{-1} (\underline{\underline{X}}^T \underline{\underline{X}})}_{\text{I}} \underline{\underline{X}}^T \bar{w}_p$$

$\lambda_p \bar{w}_p$

$$= \lambda_p (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T \cdot \bar{w}_p$$

$$= \lambda_p (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} (\underline{\underline{X}}^T \underline{\underline{X}}) \bar{a}_p = \lambda_p \bar{a}_p$$

$$(c) \quad \begin{array}{l} d=1000 \\ N=100 \end{array} \Rightarrow \underline{\underline{X}} \text{ har endast rang } 100 \text{ (99)}$$

$$\Rightarrow \text{Endast } 100 \text{ egenvärden är inte noll}$$

Snabbare!

Mindre minne!

$\Rightarrow$ Kan hitta egenvärde & egenvektorer
för de 100 mässbara genen att istället
analysera matrisen $\underline{\underline{X}}^T \underline{\underline{X}}$ som
har dimensionen 100×100

$\Leftarrow$