Subject :

Year . Month . Date .

★ Session 10 : CSP

ـ در این جلسه می‌خواهیم کاربرد دیگری از مطالب گفته شده در ماشین لرنینگ

را بررسی کنیم. مسئله ماشین لرنینگ BCI را به صورت زیر تعریف می‌کنیم:

★ Brain Comp. Interface (BCI)

مسئله ماشین لرنینگ را به این صورت تعریف می‌کنیم که قرار است از سیگنال‌های دریافت

شده از سنسورهای کارگذاشته شده روی سر متوجه جهت حرکت (راست و چپ) شویم.

برای حل این مسئله از الگوریتم CSP استفاده می‌کنیم:

Common Spatial Pattern ⟶ CSP

ـ فرض کنیم در مرحله train (آموزش) دو ماتریس مشاهدات یکی مربوط به حرکت

به سمت راست و دیگری سمت چپ دریافت می‌کنیم. ⟵ X_1^{r} , X_1^{l}

باید یک ویژگی در دو ماتریس مشاهدات بدست آوریم تا حرکت به چپ و راست را از هم

متمایز کنیم. به این منظور یک بردار (W) را از چپ در بردار مشاهدات ضرب می‌کنیم و

Batus

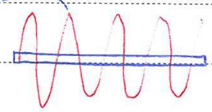
Subject :

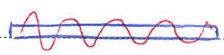
Year . Month . Date .

طوری این w را تعیین می‌کنیم که واریانس بدست آمده برای هر حرکت محدوده کاملاً متفاوتی داشته باشد.

نگاشت خطی

واریانس زیاد

$w^T X_1^r =$  $\xrightarrow{\text{variance}} \sigma_1^{2(r)}$

$w^T X_1^\ell =$  $\xrightarrow{\text{variance}} \sigma_1^{2(\ell)}$

واریانس کم

* حال تابع هدف $f(w)$ را بصورت زیر تعریف می‌کنیم

$$f(w) = \frac{\| w^T X_1^r \|^2}{\| w^T X_1^\ell \|^2} = \frac{w^T \underbrace{X_1^r X_1^{rT}}_{R_x^r} w}{w^T \underbrace{X_1^\ell X_1^{\ell T}}_{R_x^\ell} w} = \frac{w^T R_x^r \, w}{w^T R_x^\ell \, w}$$

* درصورتی که چندین مشاهده (train های دیگری) هم داشته باشیم:

$$R_x^{(r)} = \frac{1}{N_r} \sum_{n=1}^{N_r} X_n^r X_n^{rT} \quad , \quad R_x^\ell = \frac{1}{N_\ell} \sum_{n_1=1}^{N_\ell} X_n^\ell X_n^{\ell T}$$

میانگین

$$\rightarrow f(w) = \frac{w^T R_x^{(r)} w}{w^T R_x^{(\ell)} w} \quad , \quad s.t. \quad w^T w = 1$$

$$\nabla_w f = \frac{2R_x^{(r)} w \,(w^T R_x^{(\ell)} w) - 2R_x^{(\ell)} w \,(w^T R_x^{(r)} w)}{\bigcirc^2} = 0$$

Batus

Subject :

Year . Month . Date .

$$\rightarrow \cancel{2} R_x^{(r)} w \, (w^T R_x^{(l)} w) = \cancel{2} R_x^{(l)} w \, (w^T R_x^{(r)} w)$$

$f(w)$ همان

λ

$$\rightarrow R_x^{(r)} w = \frac{w^T R_x^{(r)} w}{w^T R_x^{(l)} w} \cdot R_x^{(l)} w$$

نسبت ریلی

$$\rightarrow R_x^{(l)^{-1}} R_x^{(r)} w = \lambda w$$

Generalized EVD (GEVD)

که این همان فرم مقدار و بردار ویژه برای ماتریس $R_x^{(l)^{-1}} R_x^{(r)}$ است

در محیط متلب $\Rightarrow [U, \Lambda] = eig(R_x^{(r)}, R_x^{(l)})$

$$\Lambda \Rightarrow \lambda_1 > \lambda_2 > \dots > \lambda_M$$

$$U \Rightarrow u_1 \quad u_2 \quad u_M$$

$$w_1 \quad w_M$$

در حقیقت ویژگی استخراج شده از بردار $w^T X_1$، واریانس (انرژی) آن بود. در ماشین لرنینگ بهتر است چند ویژگی را استخراج کنیم.

$w^T X_1 \xrightarrow{\text{فشرده سازی}} \sigma_1^2 \Rightarrow$ یک سری اطلاعات در فشرده سازی از دست رفته است.

Batus

Subject :

Year . Month . Date .

★نکته: در وکتور w بدست آمده، عنصر مربوط به هر سنسور را بدست می‌آوریم و کل w ها را رسم می‌کنیم. هرکدام مقدار بزرگتری داشت یعنی در واریانس‌ها موثرتر است و می‌توانیم از آن سنسور در برابر استفاده کنیم. (به جای w_i سنسور همان i سنسور را استفاده کنیم.)

★ می‌توانیم برعکس ویژگی ۱، این بار w را طوری تعریف کنیم که واریانس حرکت به چپ خیلی بزرگتر از واریانس حرکت به راست شود.

★ ویژگی ۱: تعریف w: $var(w^T X_1^{(r)}) >> var(w^T X_1^{(l)})$

★ ویژگی ۲: تعریف w: $var(w^T X_1^{(r)}) << var(w^T X_1^{(l)})$

مراحل بدست آوردن w برای ویژگی ۲ مثل ویژگی ۱ طی می‌شود با این تفاوت که تابع هدف معکوس $\frac{1}{f(w)}$ می‌شود. یعنی ماکزیمم ویژگی ۲، مینیمم ویژگی ۱ می‌شود. پس از خروجی تابع eig می‌توانیم w_M را به عنوان ویژگی ۲ استخراج کنیم.

$$[U, \Lambda] = eig(R_x^{(r)}, R_x^{(l)})$$

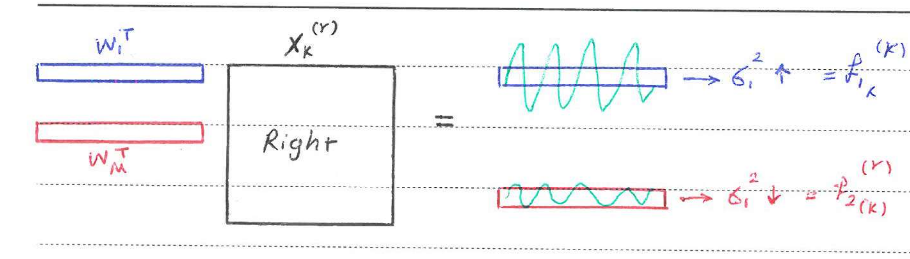
$\Lambda \rightarrow \lambda_1 > \lambda_2 > \dots > \lambda_M$

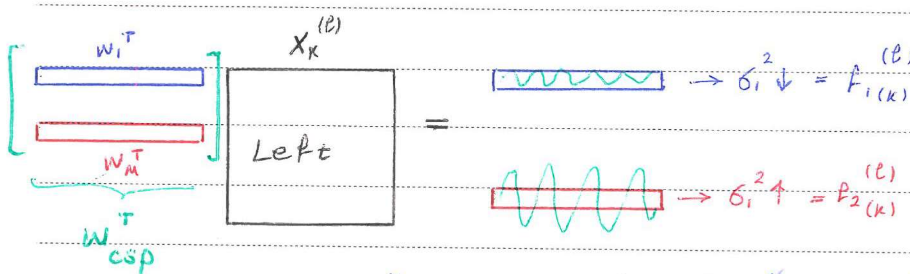
$U \rightarrow u_1, u_2, \dots, u_M$

$u_1 \rightarrow w_1$ ویژگی ۱

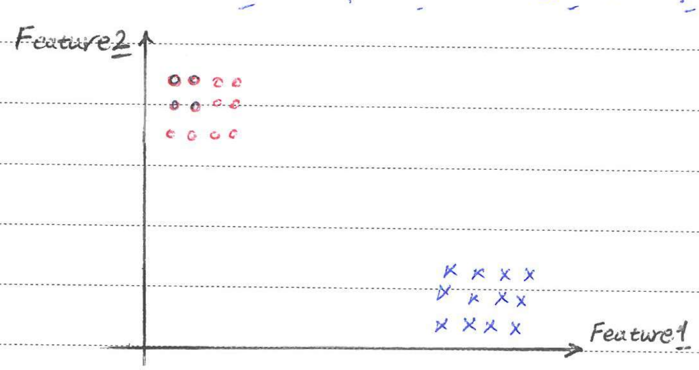
$u_M \rightarrow w_M \rightarrow$ ویژگی ۲

Batus





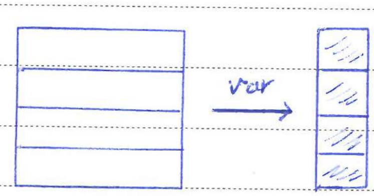
* در نتیجه فضای ویژگی (σ^2 هر ویژگی) به صورت زیر می شود:



در الگوریتم CSP تعداد زوج ویژگی انتخاب می کنیم و در W_{csp} قرار می دهیم. مثلاً:

$$W_{csp} = [\, w_1 \quad w_2 \;\vdots\; w_{M-1} \quad w_M \,]$$

$$W_{csp\,4\times T}^{T} \; X_{M\times T}^{(r)} =$$



Subject :

Year .   Month .   Date .

★ درحالت کلی اگر 2m ویژگی داشته باشیم :   2m features

$$W_{csp_{M\times 2m}} = [\, w_1, w_2, \ldots, w_m \,\vdots\, w_{M-m+1}, \ldots, w_{M-1}, w_M \,]$$

بعد از استخراج ویژگی باید داده‌ها را کلاس بندی کنیم. یعنی مرز ویژگی‌ها را مشخص کنیم.

★ طبقه بند : ★ Linear Discriminant analysis (LDA)

طبقه بند خطی دو کار می‌کند : ① ← میانگین‌ها را از هم دور می‌کند.

② ← واریانس‌ها را کم می‌کند.

← حال ابتدا رابطه میانگین‌ها و واریانس‌ها را بدست می‌آوریم :

اسکالر

$$\mu^{(1)} = \frac{1}{N_1}\sum_{t=1}^{N_1} w_{LDA}^T x^{(1)}(t) = w_{LDA}^T \underbrace{\left(\frac{1}{N_1}\sum_{t=1}^{N_1} x^{(1)}(t)\right)}_{\mu_1} = w_{LDA}^T \mu_1$$

$$\mu^{(1)} = w_{LDA}^T \mu_1 \quad , \quad \mu^{(2)} = w_{LDA}^T \mu_2$$

اسکالر

$$\sigma^{2\,(1)} = \frac{1}{N_1}\sum_{t=1}^{N_1} \left(w_{LDA}^T x^{(1)}(t) - \mu^{(1)}\right)^2$$

$$= \frac{1}{N_1}\sum_{t=1}^{N_1} \left(w_{LDA}^T \left(x^{(1)}(t) - \mu_1\right)\left(x^{(1)}(t) - \mu_1\right)^T w_{LDA}^T\right)$$

Subject :

Year .   Month .   Date .

$$\rightarrow \sigma^{2^{(1)}} = W_{LDA}^T \underbrace{\left[\frac{1}{N_1} \sum_{t=1}^{N_1} (x^{(1)}(t) - \mu_1)(x^{(1)}(t) - \mu_1)^T \right]}_{\Sigma_1} W_{LDA}$$

$$\circledast \rightarrow \Sigma_1 = \frac{1}{N_1} \sum_{t=1} (x^{(1)}(t) - \mu_1)(x^{(1)}(t) - \mu_1)^T$$

$$\sigma^{2^{(1)}} = W_{LDA}^T \Sigma_1 W_{LDA} \quad , \quad \sigma^{2^{(2)}} = W_{LDA}^T \Sigma_2 W_{LDA}$$

$$f(W_{LDA}) = \frac{(\mu^{(1)} - \mu^{(2)})}{\sigma^{2^{(1)}} + \sigma^{2^{(2)}}} = \frac{(W_{LDA}^T \mu_1 - W_{LDA}^T \mu_2)^2}{W_{LDA}^T \Sigma_1 W_{LDA} + W_{LDA}^T \Sigma_2 W_{LDA}}$$

$$= \frac{W_{LDA}^T \overset{\circledast\circledast}{(\mu_1 - \mu_2)(\mu_1 - \mu_2)^T} W_{LDA}}{W_{LDA}^T \underset{\circledast}{(\Sigma_1 + \Sigma_2)} W_{LDA}}$$

بردار ویژه متناظر با

بزرگترین مقدار ویژه

$$W_{LDA} = GEVD(\circledast, \circledast\circledast)$$

در محیط متلب

eig

$$[U, \Lambda] = eig(\circledast, \circledast\circledast) \Rightarrow \lambda_1 > \lambda_2 > \cdots > \lambda_M$$

$\lambda_1 \downarrow u_1 \downarrow W_{LDA}$; $\lambda_M \downarrow u_M$

Subject :

Year . Month . Date .

★ برای اینکه مرز بین دو گروه را مشخص کنیم باید یک آستانه (thr) تعریف کنیم. بهترین گزینه، میانگین میانگین دو گروه است، یعنی:

$$thr = \frac{1}{2}\left(\mu^{(1)} + \mu^{(2)}\right)$$

★ جلسه یازدهم:

★ کاربرد fft در متلب:

$x(t) =$