CS 4786, lec. 7 (2/19/15)

<clickers ??>

AI ~~Administrative~~: ☑

contents ☑

do first.

~~Settling bit code~~

⓪ updated hw

① Grouping on CMS. - q's on the grouping policy.

② Acad. int. { working together (just w/in group - if sthg happens, tell us)
- if you find helpful material on the internet / references: cite it, and translate to our notation; conventions.

④ (how to generate random data)
- scripts that are available
  - (2d-gaussian.R: uses package mvrnorm in package MASS <learned from searching around>)
  - in general: plotting; generating data
  - computing covar matrices, correlation matrices, eigen~~...~~ vectors, PCA, CCA : lecture scripts, some new.
  - visualizing smileys

CCA using scikit-learn: problematic!! (see email to Karthik).

① use sign for RP, note cheatsheet has typo.

- CCA: relat'n btwn "lin combo" and projection? (2 views!!)

- ~~⊕~~ transpose.

~~amounts~~ of changes.

③ TA expertise

question 0:

We've referred in the problem set to "projections" y_t.

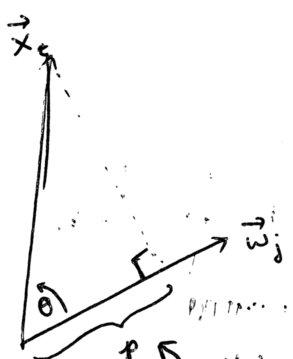
[y is k-dimensional]

Is the projection of a vector onto another vector (~~entries~~ an entry $y_t[j]$)

(a) a vector
(b) a distance (always ≥ 0)
(c) a signed distance
(d) none of the above
(e) question is ill-posed

We're getting these projections by inner products: $y_t = \begin{bmatrix} -w_1- \\ \vdots \\ -w_k- \end{bmatrix} \begin{bmatrix} x_t \end{bmatrix}$ ~~<lec 2>~~

Let's take $\vec{x}$ and project onto $\vec{w}$ a unit-vector.



$$
\begin{aligned}
x_t \cdot w_j &= \|x_t\|_2 \|w_j\|_2 \cos(\theta) \\
&= \|x_t\|_2 \cos(\theta) \\
&= \cancel{\|x_t\|_2} \frac{p}{\cancel{\|x_t\|_2}} \quad \leftarrow \text{can be negative}
\end{aligned}
$$

p ← this # (which can be negative: $\theta > 90$) is $y_t[j]$

~~(not~~

Q0 "Orientation" of vectors to project and to project on to.

Q1. Suppose you are given a matrix and data matrix

$$B = \begin{array}{c} \overbrace{\hphantom{— b_1^T —}}^{d} \\ \begin{bmatrix} — b_1^T — \\ \vdots \\ — b_k^T — \end{bmatrix} \end{array} \qquad X = \begin{array}{c} \overbrace{\hphantom{— x_1^T —}}^{d} \\ \begin{bmatrix} — \mathbf{x}_1^T — \\ \vdots \\ — \mathbf{x}_n^T — \end{bmatrix} \end{array}$$

and you want a vector y_t: $\begin{bmatrix} \\ \\ \\ \end{bmatrix}$ where $y_t[j]$ = projection of $\mathbf{x}_t$ on $\mathbf{b}_j$

(a) take t^{th} row of BX^T (and transpose it if you want "tall" vectors y_t)

(b) take t^{th} row of XB^T (")

(c) take the t^{th} column of BX^T

(d) take the t^{th} column of XB^T

(e) exactly two of the above are correct [and you know which ones]

$$BX^T: \begin{array}{c} \\ k \end{array} \begin{array}{c} d \\ \begin{bmatrix} — b_j^T — \end{bmatrix} \end{array} \begin{array}{c} \\ d \end{array} \begin{array}{c} n \\ \begin{bmatrix} | \\ \mathbf{x}_t \\ | \end{bmatrix} \end{array}$$

→ $k \times n$, suggesting we want the columns of this.

t^{th} column = inner product of $\mathbf{x}_t$ w/ each of the $\mathbf{b}_j$'s.

$$XB^T: n\left\{ \begin{array}{c} \overbrace{\hphantom{— x_t^T —}}^{d} \\ \begin{bmatrix} — \mathbf{x}_t^T — \end{bmatrix} \end{array} \right. \begin{array}{c} \\ d \end{array} \begin{array}{c} k \\ \begin{bmatrix} | \\ b_j \\ | \end{bmatrix} \end{array}$$

→ $n \times k$, suggesting we want the rows of this.

t^{th} row entries: inner products of x_t w/ each of the b_js.

<as observed in class, BX^T is the transpose of XB^T. You can directly compute from that, or use it to sanity-check your results>

- the point is that you want inner products, which you can express via matrices in various ways.

Q2: Suppose you are given a matrix and data matrix

$$C = \underbrace{\begin{bmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_k \\ | & | & & | \end{bmatrix}}_{}\Big\} d \qquad X = \underbrace{\begin{bmatrix} — & x_1^T & — \\ & \vdots & \\ — & x_n^T & — \end{bmatrix}}_{d}$$

and you want a vector y_t where $y_t[j]$ = projection of x_t on c_j.

(a) take the t^{th} row of $C^T X^T$ (and transpose to get a "tall" y_t)

(b) take the t^{th} row of XC (")

(c) take the t^{th} column of $C^T X^T$

(d) take the t^{th} column of XC

(e) exactly two of the above are correct [and you know which ones]

<skipped – left as an exercise>

CA: you had two views, given by your two diff profs:

view 1: find linear combinations of the two sets of features that have max correlation.

view 2: find projections for ~~the~~ each of the two views ~~that max correlation~~ based on common info btwn 2 views.

Are these reconciliable?
Yes: linear combination coefficients = projections (as numbers)
correlation $\equiv$ information (in a linear sense)