CPSC 340: Machine Learning and Data Mining

PCA: fit

Bonus slides

PCA Computation: SVD

- How do we fit with normalization/orthogonality/sequential-fitting?
  - It can be done with the "singular value decomposition" (SVD).
  - Take CPSC 302.

- 4 lines of Julia code:
  - mu = mean(X,1)
  - X -= repmat(mu,n,1)
  - (U,S,V) = svd(X)
  - W = V[:,1:k]'

Computing $\tilde{Z}$ is cheaper now:

$$\tilde{Z} = \tilde{X}W^T(\underbrace{WW^T}_{})^{-1} = \tilde{X}W^T$$

$$WW^T = \begin{bmatrix} - & w_1 & - \\ - & w_2 & - \\ & \vdots & \\ - & w_k & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ w_1^T & w_2^T & \cdots & w_k^T \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & & \ddots & & 0 \\ \vdots & & & & \\ 0 & \cdots & & 0 & 1 \end{bmatrix} = I$$

Making PCA Unique

- PCA implementations add constraints to make solution unique:
  - Normalization: we enforce that $||w_c|| = 1$.
  - Orthogonality: we enforce that $w_c^T w_{c'} = 0$ for all $c \neq c'$.
  - Sequential fitting: We first fit w_1 ("first principal component") giving a line.
    - Then fit w_2 given w_1 ("second principal component") giving a plane.
    - Then we fit w_3 given w_1 and w_2 ("third principal component") giving a space.
    - …
- Even with all this, the solution is only unique up to sign changes:
  - I can still replace any w_c by $-w_c$:
    - $-w_c$ is normalized, is orthogonal to the other $w_{c'}$, and spans the same space.
  - Possible fix: require that first non-zero element of each w_c is positive.
  - And this is assuming you don't have repeated singular values.
    - In that case you can rotate the repeated ones within the same plane.