PCA

11/2

maximize $\sigma_u^2 = \vec{u}^T \Sigma \vec{u}$

subject to $\|\vec{u}\| = 1 \Leftrightarrow \vec{u}^T\vec{u} = \vec{u}\cdot\vec{u} = 1 \Leftrightarrow \vec{u}\cdot\vec{u} - 1 = 0$

find a local max of a function subject to equality constraint $\Rightarrow$ use lagrange mult method

Let $\alpha \in \mathbb{R}$

set up constrained opt as unconstrained

$$J(\vec{u}) = \underbrace{\vec{u}^T \Sigma \vec{u}}_{\text{obj funct}} - \alpha \underbrace{(\vec{u}\cdot\vec{u} - 1)}_{\text{constraint}}$$

lagrange Mult

$$\frac{\partial J(u)}{\partial u} = 0$$

$$\Leftrightarrow \frac{\partial}{\partial u}\left(u^T \Sigma u - \alpha(u\cdot u - 1)\right) = 0$$

$\vdots$

vector

$\Sigma \vec{u} = \alpha \vec{u}$ $\Rightarrow$ $\vec{u}$ is an eigenvector of Σ and α is its eigenvalue

matrix scalar

$\Rightarrow$ 1st eignvector is the direction maximizing projected var!

useful tool

$$\Sigma\vec{u} = \alpha\vec{u}$$

left mult by $\vec{u}^T$ $\Leftrightarrow$ $\vec{u}^T\Sigma\vec{u} = \vec{u}^T\alpha\vec{u} = \alpha\,\vec{u}^T\vec{u} = \alpha$

1

Get 2^{nd} principal component

$\bar{D}$ data centered at mean

let u_1 be eigenvector w/ largest eigenvalue

Goal find dir $\vec{v}$:

Maximize var $\quad \sigma_v^2 = \vec{v}^T \Sigma \vec{v}$

orthogonal to u_1 $\quad \vec{v} \cdot \vec{u}_1 = 0$

unit len $\quad \vec{v} \cdot \vec{v} = 1$ — $\vec{v} \cdot \vec{v} - 1 = 0$

again use Lagrange

$$J(v) = \vec{v}^T \Sigma \vec{v} - \alpha(\vec{v} \cdot \vec{v} - 1) - \beta(\vec{v} \cdot \vec{u})$$

take deriv w.r.t. v

$$\frac{\partial J(v)}{\partial v} \overset{\max_v}{=} 2\Sigma \vec{v} - 2\alpha \vec{v} - \beta u_1 = \vec{0} \qquad (*)$$

premult by u_1^T

$u_1 \cdot v = 0$

$$2u_1^T \Sigma \vec{v} - 2\alpha \boxed{u_1^T \vec{v}} - 2\beta \boxed{u_1^T u_1} = \vec{0}$$

$\|u_1\| = 1$

$$\Leftrightarrow \quad 2u_1^T \Sigma \vec{v} - 2\beta = 0$$

as Σ is sym

$$u_1^T \Sigma \vec{v} = \vec{v}^T \Sigma u_1$$

$$\Leftrightarrow \quad \cancel{2} \vec{v}^T \boxed{\Sigma u_1} = \cancel{2}\beta$$

$\lambda_1 u_1$

$$\Leftrightarrow \quad \vec{v}^T \lambda u_1 = \beta$$

$$\parallel$$

$$\lambda (\vec{v}^T u_1) = 0 \qquad \Rightarrow \quad \beta = 0$$

$v \cdot u_1 = 0$

plug $\beta = 0$ into $*$

$$2\Sigma\vec{v} - 2\alpha\vec{v} = 0$$

$$\iff \cancel{2}\Sigma\vec{v} = \cancel{2}\alpha\vec{v}$$

$$\iff \Sigma\vec{v} = \alpha\vec{v} \Rightarrow$$ $\vec{v}$ is another eignvector $\rightarrow$ 2^{nd} principal component
α is the eignvale

Best r-dim approx

Let $2 < r \leq d$

assume we have first $j-1$ principal components w/ $j < r$

$u_1, u_2, \dots, u_{j-1}$ w/ u_ℓ as ℓ-th principal component / eigmvector

want to compute new basis vector $\vec{v}$ w/

$\vec{v} \cdot \vec{v} = 1$

$\vec{u}_i \cdot \vec{v} = 0 \quad i = 1, 2, \dots, j-1$

$\sigma_v^2 = \vec{v}^T \Sigma \vec{v}$

lagrang mult

$$J(\vec{v}) = \vec{v}\Sigma\vec{v} - \alpha(v^T v - 1) - \sum_{i=1}^{j-1} \beta_i (u_i \cdot v)$$

$$\frac{\partial J(\vec{v})}{\partial \vec{v}} = 2\Sigma\vec{v} - 2\alpha\vec{v} - \sum_{i=1}^{j-1} \beta_i u_i = \vec{0} \qquad *$$

for each $k = 1, 2, \dots, j-1$

left mult by u_k^T

$$2\underbrace{\vec{u}_k^T \Sigma \vec{v}}_{= \vec{v}^T \Sigma \vec{u}_k} - 2\alpha\overbrace{(\vec{u}_k^T \cdot v)}^{0} - \beta_k \overbrace{(u_k \cdot u_k)}^{1} - \overbrace{\sum_{\substack{i=1 \\ i \neq k}}^{j-1} \beta_i \overbrace{(u_k^T \cdot u_i)}^{0}}^{0}$$

rewrite $2v^T \Sigma \vec{u}_k - \beta_k = 0$

$$2v^T \Sigma \vec{u}_k = \beta_k$$

$$2\vec{v}^T \lambda_k u_k = 2\lambda_k \underbrace{(\vec{v}^T \cdot u_k)}_{0} = 0$$

plug back into

(A) $2\Sigma\vec{v} - 2\alpha\vec{v} = 0$

$\Leftrightarrow \cancel{2}\Sigma\vec{v} = \cancel{2}\alpha\vec{v}$

$\Leftrightarrow \Sigma\vec{v} = \alpha\vec{v} \Rightarrow \alpha = \lambda_j$ j^{th} eig val

$\vec{v} = u_j$ j^{th} eig vect