Computer Vision   Homework #1   Greg Kiar   09/10/2014

GK

1. a) $S v = \lambda v$ (1)

→ Assume $\lambda \in \mathbb{C}$, ∴

$S\bar{v} = \bar{\lambda}\bar{v}$ (2)

[citation]: Quandt, Richard E. Princeton University. "Some basic matrix theorems"

$\hat{v} S v = \hat{v}\lambda v$ (1)' $v S \hat{v} = v \hat{\lambda} \hat{v}$ (2)'

(1)' − (2)'

$\hat{v} S v - v S \hat{v} = \hat{v}\lambda v - v\hat{\lambda}\hat{v}$

∵ S is symmetric, $\hat{v} S v - v S \hat{v} = 0$, ∴

$\hat{v} v^{T}(\lambda - \hat{\lambda}) = 0 \ldots$

∵ $\hat{v}\cdot v \neq 0$, $\lambda - \bar{\lambda} = 0$, ∴ λ must be real.

b) $S v = \lambda v$

→ $S v_i = \lambda_i v_i$ (1)

$S v_j = \lambda_j v_j$ (2)

$v_j S v_i = v_j \lambda_i v_i$ (1)' $v_i S v_j = v_i \lambda_j v_j$ (2)'

$v_j S v_i - v_i S v_j = v_j \lambda_i v_i - v_i \lambda_j v_j$

∵ S is symmetric, $v_2 S v_1 - v_1 S v_2 = 0$, ∴

$v_j v_i (\lambda_j - \lambda_i) = 0 \ldots$

∵ $\lambda_j - \lambda_i \neq 0$, $v_i \perp v_j$

c) ∵ S is symmetric, and all $\lambda|_i^n \in \mathbb{R}$

$$\left.\begin{aligned} S v_1 &= \lambda_1 v_1 \\ S v_2 &= \lambda_2 v_2 \\ &\vdots \\ S v_n &= \lambda_n v_n \end{aligned}\right\}$$

∴ $v_i \perp v_j$, and λ_i can be scaled such that $|v_i| = 1$, V satisfies being orthonormal

vectors in V which are orthonormal

Also, ∵ S is $n \times n$, Eigen Value Decomposition Applies:

$$S = V \Lambda V^{-1}$$

∵ V is always invertible, V must have n vectors.

∴ V is orthonormal basis of n elements.

d) $SV = V\Lambda$, where V is the orthonormal basis of S

$$SV = \alpha_1 v_1 \lambda_1 + \alpha_2 v_2 \lambda_2 + \dots + \alpha_n v_n \lambda_n$$

[citation]: Wilson, Richard. Math, Caltech. 2010-2011. "Real symmetric matrices"

$\because v_i^T v_j = 0$ for $i \neq j$, $v_i^T v_j = 1$ for $i = j$

$$V^T S V = \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \dots + \alpha_n^2 \lambda_n$$

$\therefore$ iff $\lambda_i > 0$ for all $i = 1 \to n$, is $S \succ 0$. Also, iff $\lambda_i \geq 0$ for all $i = 1 \to n$, is $S \succeq 0$. This assumes that not all $\alpha_i = 0$.

e) $\max\limits_{\|x\|_2^2 = 1} x^T S x = \lambda_1$, $\min\limits_{\|x\|_2^2 = 1} x^T S x = \lambda_n$ $\qquad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$

$$\mathcal{L}(x, \lambda) = x^T S x + \lambda(1 - x^T x)$$

$$\frac{\delta \mathcal{L}}{\delta x} = 2Sx - 2\lambda x$$

$$0 = 2Sx - 2\lambda x$$

$$Sx = \lambda x$$

$$x^T S x = \lambda \longrightarrow \therefore \min x^T S x = \min(\lambda) = \lambda_n$$

$$\max x^T S x = \max(\lambda) = \lambda_1$$

2. a) $A = U \Sigma V^T$

$$AV = U \Sigma V^T V$$

$$AV = \Sigma U$$

$$A v_i = \sigma_i u_i$$

$$A = U \Sigma V^T$$

$$U^T A = U^T U \Sigma V^T$$

$$(U^T A = \Sigma V^T)^T$$

$$A^T U = \Sigma V$$

$$A^T u_i = \sigma_i v_i$$

b) $A = [u_r \;\; u_{r+1}^{\perp}] \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_r^T \\ v_{r+1}^{\perp T} \end{bmatrix}$

$$A = U_r \Sigma_r V_r^T$$

For u_r, spanning from $i = 1 \to r$, all values of A are included, i.e. $\text{range}(A) = \text{span}(u_r)$ not $\text{span}(u_n)$ $\because \Sigma_\alpha = 0$, $\alpha = r+1 \to n$.

c) From the eqn above, we see that $v_{r+1}^{\perp T}$ elements are always eliminated, by being multiplied by zero, $\therefore$ they make up the kernel space of A. They make up the whole kernel as they span from the range of A to n.

d) $\|A\|_F^2 = \sum_{i,j}^{n,m} a_{ij}^2$

$= \sum_{i,j}^{n,m} u_{ij}^2 \sigma_{ij}^2 v_{ij}^2$

$\because u_{ij}^2 = u_{ij} u_{ji} = I$ (same for v_{ij}^2)

$= \sum_{i,j}^{n,m} \sigma_{ij}^2$

$$\because \Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_r \end{bmatrix}$$

if $i \neq j$, $\sigma_{ij} = 0$

$i,j > r$, $\sigma_{ij} = 0$

$= \sum_{k}^{r} \sigma_k^2$

---

e) $\min_{\|x\|_2^2=1} \|Ax\|_2^2 = \min_{\|x\|_2^2=1} x^T A^T A x$

$= \min_{\|x\|_2^2=1} x^T V \Sigma U^T U \Sigma V^T x$

$= \min_{\|x\|_2^2=1} x^T (V \Lambda V) x$

↳ This is of the same form as 1·e).

∴ we can see that $\min_{\|x\|_2^2=1} \|Ax\|_2^2 = \min(\sigma) = \sigma_m$

---

3. a) $A = U_r \Sigma_r V_r^T$

$\because A A^+ A = A$, $A A^+ = I$

$\because A A^+ = I$

$U_r \Sigma_r V_r^T A^+ = I$

$U_r \Sigma_r V_r^T \underbrace{V_r \Sigma_r^{-1} U_r^T}_{A^+} = I$

$\therefore A^+ = V_r \Sigma_r^{-1} U_r^T$

---

b) $\min \|Ax - b\|_2^2 = \min (Ax-b)^T (Ax-b) = f$

$\frac{df}{dx} = 0 = 2A^T(Ax - b)$

$= 2A^T A x - 2A^T b$

$2A^T A x = 2A^T b$

$x = (A^T A)^{-1} A^T b \rightarrow$ by definition, A^+

$x = A^+ b$

∴ the unique soln is x^* when $b \neq 0$.