Prop. $$\lambda_n \le \frac{u^T A u}{u^T u} \le \lambda_1 \quad (A \text{ sym})$$

Each "$=$" holds if and only if u is the associated eigenvector.

Proof: We prove the second inequality:

$$\max_{u \neq 0} \frac{u^T A u}{u^T u} = \max_{u \neq 0} \left(\frac{u^T}{\|u\|}\right) A \left(\frac{u}{\|u\|}\right) \overset{v \overset{\text{def}}{=} \frac{u}{\|u\|}}{=\!=} \max_{\|v\|=1} v^T A v.$$

Let $A = U \underbrace{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}}_{\Lambda} U^T$ be the eigenvalue decomposition, then

$$v^T A v = \underbrace{v^T U}_{= w^T} \Lambda \underbrace{U^T v}_{\overset{\text{def}}{=} w} \overset{w \overset{\text{def}}{=} U^T v}{=} w^T \Lambda w$$

$$= \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_n w_n^2$$

where $\|w\| = \|$ $\| = \|v\| = 1$.

Then $\max\limits_{\|v\|=1} v^T A v = \max\limits_{\|v\|=1} v^T U \Lambda U^T v$

$$\overset{w = U^T v}{=} \max_{\|w\|=1} w^T \Lambda w$$

$$= \max_{\substack{\|w\|^2 = 1 \\ \text{i.e. } w_1^2 + w_2^2 + \cdots + w_n^2 = 1}} \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_n w_n^2$$

where the maximum value λ_1 is achieved if $w_1^2 = 1$ and $w_2^2 = \cdots = w_n^2 = 0$. In this case,

$w = (1, 0, \cdots, 0)^T$,

$$v = Uw = \begin{pmatrix} | & & | \\ u^{(1)} & \cdots & u^{(n)} \\ | & & | \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{matrix} | \\ u^{(1)} \\ | \end{matrix}$$

which is a normalized eigenvector

associated to λ_1.

$$u = \frac{v}{\|v\|} = v = u_1^{(1)}.$$

Proof for the first inequality is similar with "max" replaced by "min" ■

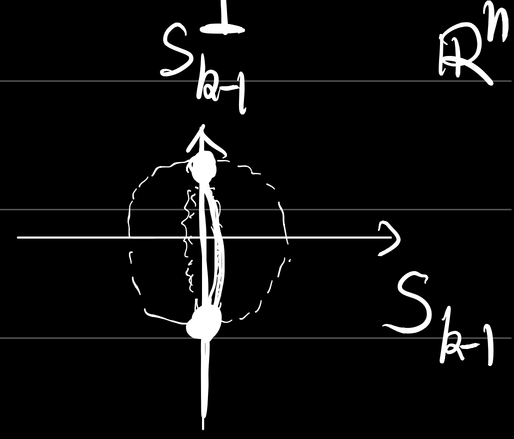
Next, we generalize the proposition to any eigenvalue. Write the eigenvalue decomp of A as

$$A = U \Lambda U^T = \begin{pmatrix} | & & | \\ u^{(1)} & \cdots & u^{(n)} \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} - & u^{(1)} & - \\ & \vdots & \\ - & u^{(n)} & - \end{pmatrix}$$

Let $S_k = \text{span}\{u^{(1)}, \ldots, u^{(k)}\}$, $S_0 = \phi$

then $S_k^{\perp} = \text{span}\{u^{(k+1)}, \ldots, u^{(n)}\}$

$$\boxed{\text{Prop.}\quad \max_{u \perp S_{k-1}} \frac{u^T A u}{u^T u} = \lambda_k \qquad k = 1, \dots, n.}$$



Proof:

$$\max_{u \perp S_{k-1}} \frac{u^T A u}{u^T u} \overset{v = \frac{u}{\|u\|}}{=\!=} \max_{\substack{\|v\|=1 \\ v \perp S_{k-1}}} v^T A v$$

$$= \max_{\substack{\|v\|=1 \\ v \perp S_{k-1}}} v^T U \Lambda U^T v$$

Set $w \overset{\text{def}}{=} U^T v = \begin{pmatrix} — u^{(1)} — \\ \vdots \\ — u^{(n)} — \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u^{(1)} \cdot v \\ \vdots \\ u^{(n)} \cdot v \end{pmatrix}$

$$\overset{v \perp S_{k-1}}{=\!=} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ u^{(k)} \cdot v \\ \vdots \\ u^{(n)} \cdot v \end{pmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} (k-1)\ 0\text{'s} \\ = w_k \\ \vdots \\ = w_n . \end{array}$$

Then $\max\limits_{\substack{\|v\|=1 \\ v \perp S_{k-1}}} v^T A v \overset{w = U^T v}{=\!=} \max\limits_{\substack{\|w\|=1 \\ w_1 = \cdots = w_{k-1} = 0}} w^T \Lambda w$

$$= \max_{\substack{\|w\|=1 \\ w_1 = \cdots = w_{k-1} = 0}} \lambda_k w_k^2 + \cdots + \lambda_n w_n^2 .$$

which achieves the maximum λ_k when $w_k = 1$ and $w_{k+1} = \cdots = w_n = 0$. ∎

Recall: (1) For 1D sample points $y_1, \cdots, y_n \in \mathbb{R}$ their sample mean is

$$\mu \overset{\text{def}}{=} \frac{1}{n} \sum_{j=1}^{n} y_j ,$$

their sample variance is

$$\frac{1}{n} \sum_{j=1}^{n} (y_j - \mu)^2$$