Instructor: Forbes Burkowski

Course: CS898

A Tutorial for the Courant-Fischer Max-Min Theorem

We need some theorems that characterize particular eigenvalues in terms of various inequalities that do not involve the corresponding eigenvectors (although the derivation of these theorems will mention these eigenvectors). We start with some preliminary observations about dimensions of vector spaces.

Suppose F is a finite-dimensional vector space with subspaces U and V. We use the following notation to specify two subspaces that are derivable from U and V:

$$U+V=\{u+v \mid u\in U, v\in V\}$$
$$U\cap V=\{x \mid x\in U \ and \ x\in V\}.$$

It is very easy to demonstrate that both $U+V$ and $U\cap V$ are also subspaces of F. Note that if we are given any basis B_U of U and B_V of V then $B_U\cup B_V$ spans $U+V$. Define:

$$B_U=\{u^{(1)},u^{(2)},\ldots,u^{(m)}\}$$
$$B_V=\{v^{(1)},v^{(2)},\ldots,v^{(n)}\}.$$

Then

$$w\in span(B_U\cup B_V)\Leftrightarrow$$
$$w=\sum_{i=1}^{m}\alpha_i u^{(i)}+\sum_{i=1}^{n}\beta_i v^{(i)}=u+v \quad \ni \quad u\in U, v\in V \Leftrightarrow w=u+v.$$

Under these assumptions we have the following theorem:

<u>Theorem 1</u>:

$$\dim(U)+\dim(V)=\dim(U+V)+\dim(U\cap V).$$

Proof:

Let $B_{UV}=\{t^{(1)},t^{(2)},\ldots,t^{(p)}\}$ be a basis for $U\cap V$. Since $B_{UV}\subseteq U$ we can, if necessary, extend this basis with m-p additional vectors to get a basis for U, call it B_U where

$$B_U=\{t^{(1)},t^{(2)},\ldots,t^{(p)},u^{(p+1)},u^{(p+2)},\ldots,u^{(m)}\}.$$

A similar argument produces a basis for V:

$$B_V = \left\{ t^{(1)}, t^{(2)}, \ldots, t^{(p)}, v^{(p+1)}, v^{(p+2)}, \ldots, v^{(n)} \right\}.$$

Now $B = B_U \bigcup B_V = \left\{ t^{(1)}, t^{(2)}, \ldots, t^{(p)}, u^{(p+1)}, u^{(p+2)}, \ldots, u^{(m)}, v^{(p+1)}, v^{(p+2)}, \ldots, v^{(n)} \right\}$ spans $U+V$ and we need to demonstrate the linear independence of B. If

$$\sum_{i=1}^{p} \alpha_i t^{(i)} + \sum_{j=p+1}^{m} \beta_j u^{(j)} + \sum_{k=p+1}^{n} \gamma_k v^{(k)} = 0 \tag{1}$$

then

$$\sum_{k=p+1}^{n} \gamma_k v^{(k)} = -\left[\sum_{i=1}^{p} \alpha_i t^{(i)} + \sum_{j=p+1}^{m} \beta_j u^{(j)} \right]$$

is a vector in U. Now

$$\sum_{k=p+1}^{n} \gamma_k v^{(k)} \in V \quad \Rightarrow \quad \sum_{k=p+1}^{n} \gamma_k v^{(k)} \in U \bigcap V$$

implying there are scalars δ_l such that

$$\sum_{k=p+1}^{n} \gamma_k v^{(k)} - \sum_{l=1}^{p} \delta_l t^{(l)} = 0.$$

Since B_V is a linearly independent set we have $\gamma_{p+1} = \gamma_{p+2} = \ldots = \gamma_n = \delta_1 = \delta_2 = \ldots = \delta_p = 0$ and so (1) reduces to

$$\sum_{i=1}^{p} \alpha_i t^{(i)} + \sum_{j=p+1}^{m} \beta_j u^{(j)} = 0.$$

But B_u is also a linear independent set so we have all α_i and β_j coefficients equal to zero as well. Consequently, the only admissible values of the coefficients in (1) are that they are all zero and so B is a linearly independent set and hence a basis for $U+V$. Consequently

$$\begin{aligned}
\dim(U+V) &= p + (m-p) + (n-p) \\
&= m + n - p \\
&= \dim(U) + \dim(V) - \dim(U \bigcap V)
\end{aligned}$$

as required.

Bounds involving Eigenvalues

Before getting to the Courant-Fischer theorem we need one more preparatory theorem.

Let A be an $n \times n$ symmetric matrix with real eigenvalues $\lambda_i \quad i = 1,2,\ldots,n.$ So, $Aq^{(i)} = \lambda_i q^{(i)}$ and $q^{(i)T} q^{(j)} = \delta_{ij} \quad i, j = 1,2,\ldots,n.$ We know that there is an orthogonal Q such that

$$A = QDQ^T \quad \text{where} \quad D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

We will assume that $\lambda_{\max} = \lambda_i \geq \lambda_2 \geq \ldots \geq \lambda_n = \lambda_{\min}$.

With A, Q, and D defined in this manner we have:

Theorem 2:

Let $W = span\{q^{(r)}, \ldots, q^{(s)}\}$ where $1 \leq r \leq s \leq n.$ Then for any $x \in W$ with $\|x\| = 1$

$$\lambda_s \leq x^T Ax \leq \lambda_r$$

Proof:

Let

$$x = \sum_{i=r}^{s} \alpha_i q^{(i)}.$$

Note that $\|x\| = 1 \quad \Rightarrow \quad \sum_{i=r}^{s} \alpha_i^2 = 1.$

Then

$$\begin{aligned} x^T Ax &= x^T \sum_{i=r}^{s} \alpha_i Aq^{(i)} \\ &= x^T \sum_{i=r}^{s} \alpha_i \lambda_i q^{(i)} \\ &= \sum_{i=r}^{s} \alpha_i \lambda_i x^T q^{(i)} \\ &= \sum_{i=r}^{s} \alpha_i \lambda_i \left(\sum_{j=r}^{s} \alpha_j q^{(j)T} \right) q^{(i)} \\ &= \sum_{i=r}^{s} \alpha_i^2 \lambda_i. \end{aligned}$$

Because $\lambda_s \le \lambda_i \le \lambda_r \quad i = r,\dots,s$ we have

$$\lambda_s = \lambda_s \sum_{i=r}^{s} \alpha_i^2 \le x^T Ax \le \lambda_r \sum_{i=r}^{s} \alpha_i^2 = \lambda_r$$

and the theorem is proven.

Since the entire set of eigenvectors $q^{(1)}, q^{(2)}, \dots, q^{(n)}$ span $\mathbb{R}^n$ we prove a corollary to this theorem due to Rayleigh and Ritz.

<u>Corollary</u>:

$$\lambda_{\min} = \min_{\|x\|=1} x^T Ax \qquad \lambda_{\max} = \max_{\|x\|=1} x^T Ax.$$

With $x \in \mathbb{R}^n$ and $\|x\| = 1$ the minimum value of $x^T Ax$ is, according to the previous theorem, bounded below by λ_n and bounded above by λ_1. Furthermore, these bounds are attainable when $x = q^{(n)}$ and $x = q^{(1)}$. This is easily observed by simple substitution into $x^T Ax$ and then simplifying while capitalizing on the fact that $q^{(n)}$ and $q^{(1)}$ are eigenvectors.

Recall that for U and V subspaces of $F = \mathbb{R}^n$ we have

$$\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V).$$

So, $U \cap V$ is not empty if $\dim(U) + \dim(V) \ge n.$ We will need this observation in the next theorem:

<u>Theorem 3</u>: (Courant-Fischer)
Let A be symmetric with eigenvalues $\lambda_{\max} = \lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_n = \lambda_{\min} \ge 0.$
Then

$$\lambda_k = \max_{V \ni \dim(V)=k} \min_{x \in V, \|x\|=1} x^T Ax.$$

Proof:
For eigenvalue λ_i let the corresponding eigenvector be $q^{(i)}$. Consider the subspace

$$U = span\left\{q^{(k)}, \dots, q^{(n)}\right\}$$
$$\Rightarrow \dim(U) = n - (k-1).$$

Let V be any subspace of $\mathbb{R}^n$ such that $\dim(V) = k.$ Since

$$\dim(U) + \dim(V) = n - (k-1) + k = n + 1$$

we know that $U \bigcap V$ is not empty and we can find $x \in U \bigcap V$ with $\|x\| = 1.$ Since $x \in U$ the previous theorem gives us $\lambda_k \geq x^T Ax.$ Also, since $x \in V$ we can write:

$$x^T Ax \geq \min_{v \in V, \|v\|=1} v^T Av$$

and so

$$\lambda_k \geq \min_{v \in V, \|v\|=1} v^T Av.$$

But this is true without any prior restriction on V aside from $\dim(V) = k$ so

$$\lambda_k \geq \max_{V \ni \dim(V)=k} \min_{x \in V, \|v\|=1} x^T Ax.$$

We next show that the inequality can be reversed. Let $W = span\{q^{(1)}, \ldots, q^{(k)}\}.$ Then for $w \in W$ and $\|w\| = 1$ we have

$$w^T Aw \geq \lambda_k = q^{(k)T} Aq^{(k)}.$$

Since this is true for all $w \in W$ (such that $\|w\| = 1$) we can write

$$\min_{w \in W, \|w\|=1} w^T Aw \geq \lambda_k.$$

For V an arbitrary subspace of $\mathbb{R}^n$ such that $\dim(V) = k$ we then have

$$\max_{V \ni \dim(V)=k} \min_{x \in V, \|x\|=1} x^T Ax \geq \min_{w \in W, w=1} w^T Aw \geq \lambda_k$$

since W is also a dimension k subspace.

This demonstrates the reverse inequality and we are finished.