MATH0328: Numerical Linear Algebra
A Second Look at Spectral Theorem and Norm

Spectral Theorem

**Theorem 1** (Spectral Theorem). Let $A \in \mathbb{R}^{n \times n}$. Then $A = A^T$ if and only if $\mathbb{R}^n$ has an orthonormal basis of eigenvectors of $A$.

These are some of the ways in which we can use this theorem.

1. If we know the eigenvalues and eigenvectors of $A$, we can recover $A$ by computing $A = O^T \Lambda O$, where $O$ is the orthogonal matrix whose columns are the orthonormalized eigenvectors of $A$ and the diagonal matrix, $\Lambda$, is such that $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. This decomposition of $A$ is very useful in other areas of mathematics as well. For example, in differential equations solutions to sets of first order linear equations ($x' = Ax$) can sometimes be written in terms of eigenvalues and eigenvectors (provided $A$ is diagonalizable).

2. $A$ is similar to $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Many properties are preserved under similarity, such as the characteristic polynomial (and hence the eigenvalues, trace, and determinant) and nilpotence. This can simplify calculations and verifications of certain properties. For example, it is especially useful when you need to compute powers of $A$:

\[
A = O^T \Lambda O \\
A^2 = O^T \Lambda O O^T \Lambda O = O^T \Lambda^2 O \\
A^k = O^T \Lambda^k O
\]

Computing $O^T \Lambda^k O$ is much simpler computationally since it only involves 2 matrix multiplications, $\Lambda^k O$ and $O^T (\Lambda^k O)$, since we know $\Lambda^k = \text{diag}(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k)$.

3. There is an orthonormal basis of $\mathbb{R}^n$ of eigenvectors of $A$, say $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$. This means that every vector in $\mathbb{R}^n$ can be decomposed as a linear combination of the eigenvectors of $A$. This is especially useful when you need to calculate something with the 2-inner product since you can use the fact that $(\vec{u}_i, \vec{u}_j)_2 = 0$ if $i \neq j$ (by orthogonality) and $(\vec{u}_i, \vec{u}_i)_2 = 1$ ($u_i$ are unit vectors). For example, since $\vec{x} = \sum_{i=1}^n x_i \vec{u}_i$, we have

\[
\|\vec{x}\|_2^2 = \left( \sum_{i=1}^n x_i \vec{u}_i, \sum_{j=1}^n x_j \vec{u}_j \right)_2 = \sum_{i,j=1}^n x_i x_j (\vec{u}_i, \vec{u}_j)_2 = \sum_{i=1}^n x_i^2
\]

4. Many matrices that arise from real-life problems are symmetric (we will see this later). Knowing specific properties about symmetric matrices is useful for both the analysis and implementation of numerical methods.
Spectral Norm of a Matrix

Theorem 2 (Spectral Norm of a Matrix). For any $A \in \mathbb{R}^n$,
\[
\|A\|_2 = \sqrt{\rho(A^T A)}
\]

Proof. First, a few observations:

- It is called the spectral norm since it is related to the spectrum (eigenvalues) of $A^T A$ (singular values of $A$).
- If $A$ is symmetric, then the calculation simplifies to $\|A\|_2 = \rho(A)$, since $A^T A = A^2$.

Highlights of the Spectral Norm proof

- $A^T A$ is symmetric and therefore the spectral theorem applies: There is an orthonormal basis of eigenvectors of $A^T A$, denoted by $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$.
- $\sigma(A^T A)$ consists of all nonnegative real numbers. This is because if $(\lambda, \vec{u})$ is an eigenpair of $A^T A$, then
\[
0 \leq \|A\vec{u}\|_2^2 = (A\vec{u}, A\vec{u})_2 = (\vec{u}, A^T A\vec{u})_2 = (\vec{u}, \lambda \vec{u})_2 = \lambda \|\vec{u}\|_2^2 = \lambda.
\]
- Any vector in $\mathbb{R}^n$ can be written as a linear combination of the eigenvectors of $A^T A$: $\vec{x} = \sum_{i=1}^{n} x_i \vec{u}_i$.

Using orthogonality of this basis, along with bilinearity of the inner product, definition of eigenvectors, and the fact that $\vec{u}_i$ are unit vectors, it is possible to simplify the following:
\[
\|Ax\|_2^2 = \sum_{i=1}^{n} \lambda_i x_i^2.
\]
- If $(\lambda_{max}, \vec{u}_{max})$ is the maximal eigenpair of $A^T A$ (so $\lambda_{max} = \rho(A^T A)$), since all eigenvalues are nonnegative, then for each $i$, $\lambda_i \leq \lambda_{max}$ and so
\[
\|Ax\|_2^2 = \sum_{i=1}^{n} \lambda_i x_i^2 \leq \sum_{i=1}^{n} \lambda_{max} x_i^2 = \lambda_{max} \sum_{i=1}^{n} x_i^2 = \lambda_{max} \|\vec{x}\|_2^2.
\]

Taking the square root of both sides (recall the square root function is increasing, so preserves the inequality), we have
\[
\|Ax\|_2 \leq \sqrt{\lambda_{max}} \|x\|_2.
\]
- Taking the maximum over all unit vectors, we have
\[
\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2 \leq \max_{\|x\|_2 = 1} \sqrt{\lambda_{max}} \|x\|_2 = \sqrt{\lambda_{max}}
\]
- This maximum value is attained for $\vec{x} = \vec{u}_{max}$ since:
\[
\|A\vec{u}_{max}\|_2^2 = (A\vec{u}_{max}, A\vec{u}_{max})_2
= (\vec{u}_{max}, A^T A\vec{u}_{max})_2
= (\vec{u}_{max}, \lambda_{max} \vec{u}_{max})_2
= \lambda_{max} \|\vec{u}_{max}\|_2^2
= \lambda_{max}
\]

Therefore $\|A\vec{u}_{max}\|_2 = \sqrt{\lambda_{max}}$, and so equality holds.