Linear Algebra I
Summary of Lectures:
Eigenvalues and Eigenvectors

Dr Nicholas Sedlmayr

1. Definition 6.1: Let $\mathbf{A}$ be an $n \times n$ matrix over a field F. Then a column vector $\mathbf{x} \in F^n$ is called an eigenvector of $\mathbf{A}$, with eigenvalue $\lambda \in F$, if $\mathbf{x} \neq 0$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

2. Theorem 6.2: A scalar λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$ iff the matrix $\mathbf{A} - \lambda\mathbb{I}_n$ has nullity $n(\mathbf{A} - \lambda\mathbb{I}_n) > 0$.

3. Theorem 6.3: Suppose λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$. Then the eigenvectors of $\mathbf{A}$ having eigenvalue λ are the non-zero vectors in $\ker(\mathbf{A} - \lambda\mathbb{I}_n) = \{\mathbf{x} : (\mathbf{A} - \lambda\mathbb{I}_n)\mathbf{x} = 0\}$.

4. Theorem 6.4: Every $n \times n$ matrix $\mathbf{A}$ over $F = \mathbb{R}$ or $F = \mathbb{C}$ has an eigenvalue λ in $\mathbb{C}$ and an eigenvector $\mathbf{x} \in \mathbb{C}^n$ with eigenvalue λ.

5. Definition 6.5: If $f : V \to V$ is a linear map, where V is a vector space over a field F, and $0 \neq v \in V$ with $f(v) = \lambda v$ for some $\lambda \in F$, then v is an eigenvector of f, with eigenvalue λ.

6. Proposition 6.6: If $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{P}$ are $n \times n$ matrices related by $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ then $\mathbf{B}$ and $\mathbf{A}$ have the same eigenvalues.

7. Lemma 6.7: Suppose that $f : V \to V$ is a linear transformation of an n-dimensional vector space V over a field F. If f has nullity of at least one then there is a basis $v_1, v_2, \ldots v_n$ such that
$$f(v_j) \in \operatorname{span}(v_1, v_2, \ldots v_{n-1}) \quad \forall j = 1, \ldots n\,.$$

8. Proposition 6.8: Let $V = \mathbb{C}^n$ be the n-dimensional vector space over $\mathbb{C}$, and suppose f is a linear transformation from V to V. Then there is a basis of V such that, with respect to this basis, the matrix of f is upper triangular.

9. Proposition 6.9: If $\mathbf{A}$ is an upper triangular matrix then the diagonal entries in $\mathbf{A}$ are precisely the eigenvalues of $\mathbf{A}$.

10. Theorem 6.10: If $\mathbf{A}$ is any upper triangular $n \times n$ matrix with entries from $\mathbb{R}$ or $\mathbb{C}$ and $\lambda_1, \lambda_2, \ldots \lambda_n$ are the diagonal entries of $\mathbf{A}$, including repetitions, then the matrix
$$(\mathbf{A} - \lambda_1\mathbb{I})(\mathbf{A} - \lambda_2\mathbb{I}) \ldots (\mathbf{A} - \lambda_n\mathbb{I})\,,$$
is the zero matrix.

11. Lemma 6.11: If $\mathbf{A}$ is an upper triangular matrix with eigenvalue λ then $\det[\mathbf{A} - \lambda\mathbb{I}] = 0$.

12. Proposition 6.12: If $\mathbf{A}$ is an $n \times n$ matrix over $\mathbb{R}$ or $\mathbb{C}$ with eigenvalue $\lambda \in \mathbb{C}$ then $\det[\mathbf{A} - \lambda\mathbb{I}] = 0$.

13. Let $\mathbf{Y}(x) = (y_1(x), y_2(x), \dots y_n(x)^T)$ and $\mathbf{Y}'(x) = \mathbf{A}\mathbf{Y}(x)$ be a system of first order linear differential equations with constant coefficients given by $\mathbf{A} = (a_{ij})$ with $a_{ij} \in \mathbb{C}$. A general solution is given by

$$\mathbf{Y}(x) = \sum_{i=1}^{r} b_i \, \mathrm{e}^{\lambda_i x} \, \mathbf{Y}_i$$

where $b_i \in \mathbb{C}$ and λ_i are the r eigenvalues of $\mathbf{A}$ with corresponding eigenvectors $\mathbf{Y}_i$. Higher order equations (with constant coefficients) can also be solved in this way by introducing new functions. For example consider $\mathbf{Y}''(x) = \mathbf{A}_1\mathbf{Y}(x) + \mathbf{A}_2\mathbf{Y}'(x)$. Let $\mathbf{Y}_2(x) = \mathbf{Y}'(x)$ and $\mathbf{Y}_1(x) = \mathbf{Y}(x)$ then we have

$$\begin{pmatrix} \mathbf{Y}_1'(x) \\ \mathbf{Y}_2'(x) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbb{I}_n \\ \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} \mathbf{Y}_1(x) \\ \mathbf{Y}_2(x) \end{pmatrix} ,$$

which can be solved in the same way as before. This generalizes in the obvious way to higher order differential equations.

14. Theorem 6.13: If $\lambda_1, \lambda_2, \dots \lambda_r$ are distinct eigenvalues of an $n \times n$ matrix $\mathbf{A}$, with $r \leq n$, with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_r$ then $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_r$ are linearly independent.

15. Definition 6.14: An $n \times n$ matrix $\mathbf{A}$ with entries from $\mathbb{R}$ or $\mathbb{C}$ is said to be diagonalizable if there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ with $\mathbf{B}$ being diagonal. $\mathbf{P}$ is said to diagonalize $\mathbf{A}$.

16. Theorem 6.15: An $n \times n$ matrix $\mathbf{A}$ with entries from $\mathbb{R}$ or $\mathbb{C}$ is diagonalizable iff $\mathbf{A}$ has n linearly independent eigenvectors.

17. In relation to theorem 6.15 we note the following:

    (a) The diagonal entries of $\mathbf{B}$ are the eigenvalues of $\mathbf{B}$ and $\mathbf{A}$.

    (b) The column vectors of a diagonalizing matrix $\mathbf{P}$ are the eigenvectors of $\mathbf{A}$.

    (c) $\mathbf{P}$, and therefore $\mathbf{B}$ is not unique given $\mathbf{A}$. Any reordering of the eigenvectors in $\mathbf{P}$ would work.

    (d) If $\mathbf{A}$ is diagonalizable then $\mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}$ and $\mathbf{A}^k = \mathbf{P}\mathbf{B}^k\mathbf{P}^{-1}$ with $\mathbf{B}$ diagonal, giving a simple expression for the powers of the matrix $\mathbf{A}$.

18. Definition 6.16: Let V be a finite dimensional inner product space, and suppose that f is a linear transformation $f : V \to V$. f is called orthogonal (if V is an inner product space over $\mathbb{R}$) or unitary (if V is an inner product space over $\mathbb{C}$) if

$$\langle u|v\rangle = \langle f(u)|f(v)\rangle , \quad \forall u, v \in V .$$

19. Proposition 6.17: Let V be a finite dimensional inner product space, and suppose $f : V \to V$ is a linear transformation. Let $e_1, e_2, \dots e_n$ be an orthonormal basis of V. Then f preserves the inner product $\langle u|v\rangle$ on V, i.e. is orthogonal or unitary, iff $f(e_1), f(e_2), \dots f(e_n)$ is orthonormal.