Linear Algebra 2017
Summary of Lectures - The Bare Bones

Dr Nicholas Sedlmayr
ndsedlmayr@gmail.com
Department of Physics and Medical Engineering
Rzeszów University of Technology
al. Powstańców Warszawy 6, 35-959 Rzeszów, Poland

Contents

1 Matrices

- Definition of an $n \times m$ matrix, $\mathbf{A} = (a_{ij})$:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

  with n rows and m columns. Addition of matrices $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$.
  - Associativity $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$,
  - commutativity $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$, and
  - existence of a zero matrix $(\mathbf{0})$ for addition: $\mathbf{A} + \mathbf{0} = \mathbf{A}$.
- Definition of multiplication of a matrix by a scalar: $\lambda\mathbf{A} = (\lambda a_{ij})$.
- Definition: Matrix multiplication of an $n \times m$ matrix $\mathbf{A}$ and an $m \times k$ matrix $\mathbf{B}$ is an $n \times k$ matrix $\mathbf{C} = \mathbf{AB} = (c_{ij})$ where $c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$.
  - Associativity, existence of a zero matrix $(\mathbf{0})$ and an identity matrix $\mathbb{I}$, distributivity.
  - No commutativity! $\mathbf{AB} \neq \mathbf{BA}$ in general.
- A matrix $\mathbf{A}$ can have a right inverse $\mathbf{AB} = \mathbb{I}$ and a left inverse $\mathbf{CA} = \mathbb{I}$.
  - Proposition: If a square matrix, $\mathbf{A}$, has either a left or right inverse ($\mathbf{AB} = \mathbb{I}$ or $\mathbf{AC} = \mathbb{I}$) then they have a unique inverse from both the left and right. I.e. $\mathbf{B} = \mathbf{C}$.
  - Proposition: If $\mathbf{A}$ and $\mathbf{B}$ are invertible square matrices then $\mathbf{AB}$ is also invertible and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
- The transpose of an $n \times m$ matrix is written as $\mathbf{A}^T$, which is an $m \times n$ matrix found by transposing the rows and columns of $\mathbf{A}$.
  - Proposition: For two $n \times n$ matrices $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ and $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.
- Elementary row operations perform simple operations on the rows of an $n \times m$ matrix $\mathbf{A}$ and can be written as an $n \times n$ matrix $\mathbf{R}$ with the operation preformed by the multiplication $\mathbf{RA}$. ρ_i is used to refer to row i. There are three of them:
  - $\rho_j := \rho_j + \lambda\rho_i$, add λ copies of row i to row j;
  - $\rho_i := \lambda\rho_i$, multiple row i by λ with $\lambda \neq 0$; and
  - $\mathrm{swap}(\rho_i, \rho_j)$, swap rows i and j.
- Echelon form: A matrix where each row starts with a sequence of zeros, and the number of zeros in this sequences increases from row to row from top to bottom until the final row is reached or all remaining rows are composed entirely of zeros is said to be in echelon form.

- All matrices can be put into echelon form using elementary row operations.

- Rank: If $\mathbf{A}$ can be converted to the echelon form matrix $\mathbf{B}$ and $\mathbf{B}$ has k non-zero rows then the rank of $\mathbf{A}$ is $\operatorname{rk}\mathbf{A} = k$.

- An $n \times k$ matrix $\mathbf{A}$ with $n \leq k$ and $\operatorname{rk}\mathbf{A} = n$ can be converted to a matrix $\mathbf{B}$ where the left $n \times n$ block is the identity matrix using elementary row operations.

- Augmented matrix: If we have an $n \times m$ matrix $\mathbf{A}$ and an $n \times k$ matrix $\mathbf{A}$ the augmented matrix $(\mathbf{A}|\mathbf{B})$ is the $n \times (m+k)$ matrix created by writing $\mathbf{A}$ and $\mathbf{B}$ next to each other.

- Calculating the inverse: For an $n \times n$ matrix $\mathbf{A}$ of rank n we can calculate the inverse using the augmented matrix $(\mathbf{A}, \mathbb{I})$. Apply row operations, $\mathbf{R}$, such that $\mathbf{R}(\mathbf{A}, \mathbb{I}) = (\mathbb{I}, \mathbf{B})$, clearly $\mathbf{B} = \mathbf{R}$ is the left inverse of $\mathbf{A}$: $\mathbf{BA} = \mathbb{I}$.

- Solving systems of linear equations. A system of linear equations on the variables $x_1, x_2, \ldots x_n$ can be written as $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is a $k \times n$ matrix, $\mathbf{x} = (x_1, x_2, \ldots x_n)^T$ and $\mathbf{b}$ is a $k \times 1$ column vector. By putting the augmented matrix $(\mathbf{A}|\mathbf{b})$ into echelon form one can read off the solution.

  - Proposition: An equation $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is a $k \times n$ matrix, $\mathbf{x} = (x_1, x_2, \ldots x_n)^T$ and $\mathbf{b}$ is a $k \times 1$ column vector has at least one solution iff $\operatorname{rk}\mathbf{A} = \operatorname{rk}(\mathbf{A}|\mathbf{b})$. It has exactly one solution if $\operatorname{rk}\mathbf{A} = n$

- Trace: This operation is defined only for square matrices. A square $n \times n$ matrix $\mathbf{A} = (a_{ij})$ has a trace $\operatorname{tr}\mathbf{A} = \sum_{i=1}^{n} a_{ii}$. I.e. it is the sum over all diagonal entries of the matrix. It has the following properties:

  - For two square $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, $\operatorname{tr}\mathbf{AB} = \operatorname{tr}\mathbf{BA}$.
  - If $\mathbf{P}$ is a square $n \times n$ invertible matrix and $\mathbf{A}$ is a square $n \times n$ matrix then $\operatorname{tr}\mathbf{P}^{-1}\mathbf{AP} = \operatorname{tr}\mathbf{A}$.
  - $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr}\mathbf{A} + \operatorname{tr}\mathbf{B}$.
  - $\operatorname{tr}\mathbf{A}^T = \operatorname{tr}\mathbf{A}$.
  - For a scalar λ then $\operatorname{tr}\lambda\mathbf{A} = \lambda\operatorname{tr}\mathbf{A}$.

- Determinant: This operation is defined only for square matrices and we will define it via "expansion by the first row". For a 1×1 matrix $\mathbf{A} = (a_1 1)$ we have $\det\mathbf{A} = |\mathbf{A}| = a_{11}$. For an $n \times n$ matrix $\mathbf{A} = (a_{ij})$ it is

$$\begin{aligned}
\det\mathbf{A} &= \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \\
&= a_{11} \begin{vmatrix} a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n3} & \ldots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \ldots & a_{nn} \end{vmatrix} \\
&\quad + \ldots + (-1)^{n-1} a_{1n} \begin{vmatrix} a_{21} & a_{22} & \ldots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{n,n-1} \end{vmatrix}.
\end{aligned}$$

- Proposition: For two $n \times n$ square matrices $\mathbf{A}$ and $\mathbf{B}$, $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$.
- Proposition: For an $n \times n$ matrix $\mathbf{A}$ the following are equivalent:
  (a) $\mathbf{A}^{-1}$ exists,
  (b) $\det \mathbf{A} \neq 0$, and
  (c) $\operatorname{rk} \mathbf{A} = n$.
- Proposition: For an $n \times n$ matrix $\mathbf{A}$:
  (a) $\det(\mathbf{A}^T) = \det \mathbf{A}$,
  (b) If $\mathbf{A}$ is invertible $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$,
  (c) $\det \mathbb{I} = 1$, and
  (d) if $\mathbf{A}$ has a row (or column) entirely composed of zeros $\det \mathbf{A} = 0$.
- Proposition: The determinant of an upper triangular matrix $\mathbf{A}$ is equal to the product of its diagonal entries.
- Proposition: If we swap any two rows in a determinant then the determinant changes sign. It follows that if a matrix has two identical rows then its determinant is zero. (This is equally true for a matrix with two identical columns.)
- Proposition: A determinant can be expanded along any row (or column). The sign associated with any entry a_{ij} is $(-1)^{i+j}$.

2 Vector Spaces

- Definition: A vector space. A vector space V over the real numbers $\mathbb{R}$ is a set containing:

  - a special zero vector $\mathbf{0}$;
  - an operation of addition of two vectors $\mathbf{u}+\mathbf{v} \in V$, for $\mathbf{u}, \mathbf{v} \in V$; and
  - multiplication of a vector V with a number $\lambda \in \mathbb{R}$ with $\lambda\mathbf{v} \in V$.

  The vector space must be closed under both of these operations and must satisfy the following laws for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\lambda, \mu \in \mathbb{R}$:

  (1) associativity $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$;
  (2) commutativity $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$;
  (3) $\mathbf{u}+\mathbf{0} = \mathbf{u}$;
  (4) $\mathbf{v}+(-1)\mathbf{v} = \mathbf{0}$;
  (5) $\lambda(\mu\mathbf{u}) = (\lambda\mu)\mathbf{v}$;
  (6) distributivity $\lambda(\mathbf{u}+\mathbf{v}) = \lambda\mathbf{u}+\lambda\mathbf{v}$; and
  (7) distributivity $(\lambda+\mu)\mathbf{u} = \lambda\mathbf{u}+\mu\mathbf{u}$.

- Proposition: For all $\mathbf{v} \in V$ and for all $\lambda \in \mathbb{R}$:

  (a) $\mathbf{v} = 1\mathbf{v}$;
  (b) $0\mathbf{v} = \mathbf{0}$; and
  (c) $\lambda\mathbf{0} = \mathbf{0}$.

- Definition: Given a vector space V, a subspace of V is a subset $W \subset V$ which contains the zero vector of V and is closed under the operations of addition and scalar multiplication.

- Lemma: Let $W \subset V$ be nonempty, where V is a vector space. Then W is a subspace of V if and only if $\mathbf{v}+\lambda\mathbf{u} \in W$ for each $\mathbf{v}, \mathbf{u} \in W$ and each scalar λ.

- Definition: Given a vector space V, and given a subset of V: $A = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots \mathbf{u}_n\}$, then

  $$W = \{\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \lambda_3\mathbf{u}_3 + \ldots \lambda_n\mathbf{u}_n : \lambda_1, \lambda_2, \ldots \lambda_n \in \mathbb{R}\}$$

  is the subspace of V spanned by A. The elements of W are called linear combinations of vectors from A and the subspace W is denoted as span A.

- Definition: A set $A \subset V$ of vectors in a vector space V over $\mathbb{R}$ is linearly dependent if there are $n \in \mathbb{N}$ vectors $a_1, a_2, \ldots a_n$ and scalars $\lambda_1, \lambda_2, \ldots \lambda_n$ not all zero such that

  $$\lambda_1 a_1 + \lambda_2 a_2 + \ldots \lambda_n a_n = 0\,.$$

  Otherwise A is linearly independent.

- For a finite set $A = \{a_1, a_2, \dots a_n\}$ it is linearly independent iff $\forall$ scalars $\lambda_1, \lambda_2, \dots \lambda_n \in F$

$$\lambda_1 a_1 + \lambda_2 a_2 + \dots \lambda_n a_n = 0 \Rightarrow \lambda_1 = \lambda_2 = \dots \lambda_n = 0.$$

- If A is infinite it is linearly independent iff every subset of A is linearly independent.
- By convention the empty set is linearly independent.

- Proposition: Suppose $\mathbf{A} = \{a_1, a_2, \dots a_n\} \subset V$ is linearly independent, where V is a vector space over $\mathbb{R}$. Suppose also that $v \in V$ and there are scalars $\lambda_1, \dots \lambda_n$ and $\mu_1, \dots \mu_n$ such that

$$v = \lambda_1 a_1 + \lambda_2 a_2 + \dots \lambda_n a_n$$

and

$$v = \mu_1 a_1 + \mu_2 a_2 + \dots \mu_n a_n$$

then $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$,$\dots \lambda_n = \mu_n$.

- Definition: A basis of a vector space V is a linearly independent set $B \subset V$ which spans V.

- Theorem: Suppose S and B are both bases of a vector space V over $\mathbb{R}$ and. Then A and B have the same number of elements.

- Definition: The number of elements of a basis of a vector space V over $\mathbb{R}$ is called the dimension of V and is written as $\dim V$.

- Corollary: If V is a vector space over $\mathbb{R}$ and $U \subset V$ is a subspace of V then $\dim U \leq \dim V$. If, additionally, $\dim V$ is finite and $U \neq V$ then $\dim U < \dim V$.

- Corollary: Suppose that V is a vector space over $\mathbb{R}$, $\dim V$ is finite, and $U \subset V$ is a subspace of V with $\dim U = \dim V$, then $U = V$.

- The coordinates of a vector $v \in V$, with V a vector space over $\mathbb{R}$, with respect to an *ordered* basis $B = \{v_1, v_2, \dots v_n\}$ are $(\lambda_1, \lambda_2, \dots, v_n)^T$ where

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots \lambda_n v_n \,.$$

3 Inner Product Spaces

- Definition: If V is a vector space over $\mathbb{R}$, then an inner product on V is a map $(\langle | \rangle)$ from $V \times V$ to $\mathbb{R}$ with the following properties:
  (a) Symmetry: $\langle v|w\rangle = \langle w|v\rangle \ \forall v, w \in V$.
  (b) Linearity: $\langle u|\lambda v+\mu w\rangle = \lambda\langle u|v\rangle+\mu\langle u|w\rangle \ \forall u, v, w \in V$ and $\forall \lambda, \mu \in \mathbb{R}$.
  (c) Positive definiteness:
    (i) $\langle v|v\rangle \geq 0 \ \forall v \in V$, and
    (ii) $\langle v|v\rangle = 0$ iff $v = 0$.

  As the inner product is linear with respect to both variables it is sometimes called bilinear.

- Definition: A finite dimensional vector space over $\mathbb{R}$ with an inner product defined is called a Euclidean space.

- Definition: The norm (or length) of a vector v is written as $||v|$ and defined by
$$||v|| = \sqrt{\langle v|v\rangle},$$
  the positive square root of the inner product of v with itself. The distance between two vectors v and w is written as $d(v, w)$ and is $d(v, w) = ||v-w||$.

- Proposition: $\forall v \in V$, where V is a Euclidean space, and $\forall \lambda \in \mathbb{R}$ then $||\lambda v|| = |\lambda| \cdot ||v||$.

- Proposition: The "Cauchy-Schwarz inequality" says that $\forall v, w \in V$, where V is a Euclidean space, then
$$|\langle v|w\rangle| \leq ||v|| \cdot ||w|| \, .$$

- Proposition: The "triangle inequality" says that $\forall v, w \in V$, where V is a Euclidean space, then
$$||v + w|| \leq ||v|| + ||w|| \, .$$

- Definition: If V is a Euclidean space, and $v, w \in V$, then v and w are said to be orthogonal if $\langle v|w\rangle = 0$. If both v and w are nonzero, then the angle between v and w is defined to be θ, $0 \leq \theta \leq \pi$ and
$$\cos\theta = \frac{\langle v|w\rangle}{||v|| \cdot ||w||} \, .$$

- Definition: A bilinear form on a real vector space V is a map $F : V \times V \to \mathbb{R}$ which $\forall u, v, w \in V$ and $\forall \alpha, \beta \in \mathbb{R}$ satisfies
  (a) $\langle \alpha u + \beta v|w\rangle = \alpha\langle u|w\rangle + \beta\langle v|w\rangle$, and
  (b) $\langle u|\alpha v + \beta w\rangle = \alpha\langle u|v\rangle + \beta\langle u|w\rangle$.

- Definition: A bilinear form on a real vector space V is symmetric if
  (c) $F(u, v) = F(v, u) \ \forall u, v \in V$.

- Definition: The matrix $\mathbf{A} = (a_{ij})$ with $a_{ij} = F(e_i, e_j)$ is called the 'matrix of the bilinear form F with respect to the ordered basis $e_1, e_2, \ldots e_n$ of V'. If F is symmetric then $\mathbf{B}$ is symmetric.

- Proposition: Suppose V is a real vector space with ordered basis $e_1, e_2, \ldots e_n$ and F is a bilinear form defined on V, with matrix $\mathbf{A}$ with respect to this basis. Then for any vectors $v, w \in V$ and their corresponding coordinate forms $\mathbf{v} = (v_1, v_2, \ldots v_n)^T$ and $\mathbf{w} = (w_1, w_2, \ldots w_n)^T$ with respect to the same basis we have
$$F(v, w) = \mathbf{v}^T \mathbf{A} \mathbf{w} \, .$$

- The base change matrix from a basis $e_1, e_2, \ldots e_n$ to $f_1, f_2, \ldots f_n$ is $\mathbf{P} = (p_{ij})$ where $f_i = \sum_{k=1}^{n} p_{ki} e_k$.

- Proposition: (The base change formula) Given two ordered bases of a Euclidean space V, $e_1, e_2, \ldots e_n$ and $f_1, f_2, \ldots f_n$ related by the base change matrix $\mathbf{P}$ from basis $e_1, e_2, \ldots e_n$ to $f_1, f_2, \ldots f_n$, suppose $\mathbf{A}$ and $\mathbf{B}$ are the matrices of the inner product with respect to $e_1, e_2, \ldots e_n$ and $f_1, f_2, \ldots f_n$. Then $\mathbf{B} = \mathbf{P}^T \mathbf{A} \mathbf{P}$.

4 Orthogonal Bases

- Definition: Two vectors v and w in an inner product space are orthogonal if $\langle v|w\rangle$=0. The set of vectors $\{v_1, v_2, \ldots\}$ is said to be orthogonal, and the vectors $v_1, v_2 \ldots$ in the set are said to be mutually orthogonal if each pair of distinct vectors v_i, v_l with $i \neq l$ are said to be an orthogonal pair, $\langle v_i|v_l\rangle = 0$.

- Definition: A set $\{w_1, w_2, \ldots\}$ of vectors in an inner product space is said to be orthonormal if $\langle w_i|w_j\rangle = \delta_{ij}$. If the orthonormal set is a basis then it is called an orthonormal basis.

- Proposition: If V is an inner product space over $\mathbb{R}$ or $\mathbb{C}$, $v_1, v_2, \ldots v_n \in V$, $v_i \neq 0\ \forall i = 1 \ldots n$, and the v_i are mutually orthogonal then $\{v_1, v_2, \ldots v_n\}$ is a linearly independent set.

5 Eigenvalues

- Definition: Let $\mathbf{A}$ be an $n \times n$ matrix. Then a column vector $\mathbf{x} \in \mathbb{R}^n$ is called an eigenvector of $\mathbf{A}$, with eigenvalue $\lambda \in \mathbb{R}$, if $\mathbf{x} \neq 0$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.
- Proposition: 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.
- Proposition: If $\mathbf{A}$ is an $n \times n$ matrix over $\mathbb{R}$ with eigenvalue $\lambda \in \mathbb{C}$ then $\det[\mathbf{A} - \lambda\mathbb{I}] = 0$.