LINEAR ALGEBRA

ALLEN ZHU

1. Linear Equations

Definition 1. The *linear combination* of vectors $\mathbf{v}_i$ with weights c_i is $\sum c_i \mathbf{v}_i$.
Span$\{\mathbf{v}_i\}$ is the set of all linear combinations of $\mathbf{v}_i$.

Definition 2. A *linear equation* has the form $\sum a_i x_i = b$. A *system of linear equations* is a collection of one or more linear equations involving the same variables x_i, and can be represented as $A\mathbf{x} = \mathbf{b}$. A system is *consistent* iff it has at least one solution $\mathbf{x}$. The set of all possible solutions is called the *solution set*, and two linear systems are *equivalent* if they share the same solution set.
$A\mathbf{x}$ is a lincomb of the columns of A. The span of the columns is the set of all consistent b.

Definition 3. The *Gaussian elimination* algorithm solves systems of linear equations systematically by applying a series of *elementary row operations* to a matrix:

(1) *Interchange*: swap two rows
(2) *Scaling*: multiply all entries of a row by a nonzero constant
(3) *Replacement*: add a multiple of one row to another

Two matrices are *row equivalent* if there a series of EROs can transform one into the other.

Remark 1. Row equivalent *augmented matrices* represent equivalent systems.

Definition 4. The forward phase of Gaussian elimination simplifies a matrix to *row echelon form*, which satisfies the following conditions:

(1) Each pivot occurs to the right of the one above it
(2) All entries below a pivot are zeros
(3) All zero rows are below any nonzero rows

The backwards phase produces the unique *reduced row echelon form*, which also satisfies:

(1) Each leading entry is 1
(2) Each leading 1 is the only nonzero entry in its column

The *pivots* for an echelon matrix are the *leading entries*, the leftmost nonzero entries of the nonzero rows. *Basic variables* correspond to pivot columns, while the rest are *free variables*.

Definition 5. A *homogeneous* system is of the form $A\mathbf{x} = \mathbf{0}$.
Vectors $\mathbf{v}_i$ are *linearly independent* iff $V\mathbf{c} = \mathbf{0}$ has only the trivial solution.

Theorem 1. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution iff there is a pivot in every column or iff the columns of a matrix A are linearly independent.

Theorem 2. The solution set of $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x_p} + \mathbf{x_h}$, where $\mathbf{x_p}$ is a solution and $\mathbf{x_h}$ is a solution of $A\mathbf{x} = \mathbf{0}$. This is an affine space, a plane translated from the origin.

Proof. $A(\mathbf{x_p} + \mathbf{x_h}) = A\mathbf{x_p} + A\mathbf{x_h} = \mathbf{b} + \mathbf{0} = \mathbf{b}$. Conversely, if there are two distinct solutions $A\mathbf{x_1} = A\mathbf{x_2} = \mathbf{b}$, then $A(\mathbf{x_1} - \mathbf{x_2}) = \mathbf{0}$. □

Remark 2. $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$ iff there is a pivot in every column, and has a solution for all $\mathbf{b}$ iff A has a pivot in every row.
A linear system has:

- A unique solution if there is a pivot in every column (no free variables)
- No solutions if the rightmost column of its augmented matrix is a pivot column
- Infinitely many solutions otherwise

2. Vector Spaces

Definition 6. A *vector space* is an abelian group of vectors (over a field of scalars) that contains $\mathbf{0}$ and is closed under linear combinations. A *subspace* is a subset that is also a vector space, i.e. a nonempty closed subset. The *basis* of a space is a linearly independent spanning subset, such as the *standard basis* $\mathbf{e}_i$ of $\mathbb{R}^n$.

Theorem 3 (Linear Dependence)**.** A list of vectors is linearly dependent iff one of the vectors is a linear combination of the ones before it, and that vector can be removed from the linearly dependent set without affecting the span.

Proof. If a vector is a linear combination of the rest, then subtracting it from both sides produces a linear dependence. If we have a linear dependence, either the first vector is zero and it is the sum of the other vectors, or one can write a linear combination by subtracting the other vectors and dividing by its coefficient. Any vector in the span is still expressible as a linear combination of the other vectors. □

Theorem 4. Any set containing $\mathbf{0}$ is linearly dependent.

Proof. A nontrivial solution is $1\mathbf{0} + 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n = \mathbf{0}$. □

Theorem 5. In a vector space, the length of any linearly independent set is less than or equal to the length of any spanning set.

Proof. Suppose u_i is linearly independent and v_i spans V. We add a vector u_1 to the spanning set to form $(u_1, v_1, \dots, v_n)$, which must be linearly dependent. Then we may remove one of the v's so that the list still spans V. We are able to remove one v each time we repeat this process since the expanded set is dependent and v_i are independent. Thus there must at least as many spanning elements as there are linearly independent elements. □

Corollary 1. A set which contains more vectors than the entries in each vector is linearly dependent, since it is longer than the basis. Thus, the columns of a matrix are linearly dependent if there are more columns than rows.

Theorem 6 (Dimension)**.** Any spanning list in a vector space can be reduced to a basis, and any linearly independent list can be extended to a basis, so a basis is a maximal LI set or a minimal spanning set. Each basis of a subspace has the same size, called the *dimension* of the subspace. The zero subspace has an empty basis and dimension zero.

Proof. For each vector in the list: if the vector is $\mathbf{0}$ or if it is in the span of those before it, remove it. The list is still a span and is now linearly independent. The other proof is analogous. Consider two bases B_1 and B_2 of a space. Since B_1 is linearly independent and B_2 is a spanning set, $|B_1| \leq |B_2|$. Vice versa, $|B_2| \leq |B_1|$, so $|B_1| = |B_2|$. □

Theorem 7. If U is a subspace of V, then $\dim U \leq \dim V$.

Proof. Any basis of U is a linearly independent set of vectors in V and can be extended to a basis of V. □

Theorem 8. Every spanning set and every linearly independent set in V with $\dim V$ elements is a basis of V.

Proof. Suppose v_i spans V. Then it can be reduced to a basis of V. But that basis must have $\dim V$ elements, so the reduction does not change the set, so v_i originally spanned V. □

3. Matrices

Definition 7. The four fundamental subspaces of a matrix A are:

- The *column space* $\text{Col}A \in \mathbb{R}^m$ is the span of the columns, containing all $\mathbf{b}$ s.t. $A\mathbf{x} = \mathbf{b}$ is consistent.
- The *nullspace* $\text{Null}A \in \mathbb{R}^n$ is the space of solutions of the homogeneous system. Each element describes a relation on the columns of A.
- The *row space* $\text{Row}A \in \mathbb{R}^n$ is the span of the rows, or $\text{Col}A^\top$.
- The *left nullspace* $\text{LNull}A \in \mathbb{R}^m$ is the nullspace of $A^\top$. The row vectors following zero rows in $\text{rref}([A|I])$ form the basis of $\text{LNull}A$.

Remark 3. EROs preserve the row space and the order of columns. EROs also preserve the null space since $(\text{rref}A)\mathbf{x} = \mathbf{0}$ is equivalent to $A\mathbf{x} = \mathbf{0}$, so matrices with the same null space have the same row space and vice versa.

Theorem 9. The nonzero rows of $\text{rref}A$ form a basis for the row space.

Theorem 10. They span the row space, and they are independent since each pivot is the only nonzero entry in its column.

Theorem 11. Each nonpivot column of a matrix is a linear combination of pivot columns to its left.

Proof. This is relation among columns, and it holds on $\text{rref}A$. □

Theorem 12. The pivot columns of a matrix form a basis for the column space.

Proof. The pivot columns span the column space since the nonpivot columns are lincombs of them. Further, the pivot columns in $\text{rref}A$ are independent since they are distinct *standard vectors*, so the original columns are also. □

Corollary 2. $\dim \text{Row } A = \#\,\text{pivots} = \dim \text{Col } A$

Theorem 13. $\text{Row}(A)$ is the orthogonal complement of $\text{Null}(A)$, and $\text{Col}(A)$ is the orthogonal complement of $\text{LNull}(A)$.

Theorem 14 (Rank-Nullity)**.** The *rank* of a $m \times n$ matrix A is the dimension of its column space, and the *nullity* of a matrix is the dimension of its nullspace. $\text{Rank } A + \text{Nullity } A = n$.

Proof. Let $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ be the basis of Null A. We can add vectors $(\mathbf{v}_1, \dots, \mathbf{v}_q)$ to form a basis of $\mathbb{R}^n$. We aim to show that $\mathbf{v}_i$ are the basis of the column space. For $\mathbf{w} \in \mathbb{R}^n$, we have

$$\begin{aligned}\mathbf{w} &= \mathbf{a}_1\mathbf{u}_1 + \cdots + \mathbf{a}_p\mathbf{u}_p + \mathbf{b}_1\mathbf{v}_1 + \cdots + \mathbf{b}_q\mathbf{v}_q \\ &= \mathbf{b}_1\mathbf{v}_1 + \cdots + \mathbf{b}_q\mathbf{v}_q\end{aligned}$$

Thus $\mathbf{v}_i$ span the column space. Next, consider weights c_i s.t.

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

Since $\mathbf{0} \in \text{Null } A$ we can write

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = d_1\mathbf{u}_1 + \cdots + d_p\mathbf{u}_p$$

But $\mathbf{v}_i$ and $\mathbf{u}_i$ are linearly independent so $c_i = 0$ and v_i are linearly independent. Thus v_i are a basis of the column space, as desired. □

Definition 8. A *linear transformation* is a map from *domain* $\mathbb{R}^m$ to *codomain* $\mathbb{R}^n$ which is additive and homogeneous, so $T(c_1\mathbf{u} + c_2\mathbf{v}) = c_1T(\mathbf{u}) + c_2T(\mathbf{v})$. The *kernel* is the space of all vectors that map to $\mathbf{0}$. It is *one-to-one* or *injective* if there is a unique input for each output. It is *onto* or *surjective* if its range equals the codomain.

Remark 4. Transformations include reflection, dilation, skew, rotation, and projection.

Theorem 15. The *standard matrix* of a linear transformation T is the matrix

$$A = [T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ \cdots\ T(\mathbf{e}_n)]$$

which uniquely satisfies $T(\mathbf{x}) = A\mathbf{x}$.

Corollary 3. The kernel and range correspond to the nullspace and rank of the matrix, so $\dim \text{kernel } T + \dim \text{range } T = n$. T is one-to-one if the kernel/nullspace is trivial and onto if the columns span the codomain.

Definition 9. If the domain of A matches the codomain of B, then their *product* AB is the transform that applies B, then A. Equivalently,

- $(AB)_{ij}$ is the *inner product* $A_{i*}B_{*j}$
- $(AB)_{*j} = AB_{*j}$, a lincomb of the columns of A
- $(AB)_{i*} = A_{i*}B$, a lincomb of the rows of B
- AB is the sum of *outer products* $\sum A_{i*}B_{*j}^\top$

Matrix multiplication satisfies the standard algebraic properties except commutativity.

Corollary 4. Since AB is a lincomb of the rows of B, $\text{rank}AB \leq \text{rank}B$, and similarly $\text{rank}AB \leq \text{rank}A$.

Definition 10. The *left inverse* of matrix A is B s.t. $BA = I$, and *right inverse* is analogous.

Remark 5. The left inverse can be found be reducing $[A|I]$ to $[I|B]$, since $MA = I$ and $B = MI = I = I^{-1}$. The left inverse exists iff there is a pivot in every row, and the right inverse exists iff there is a pivot in every column (columns are independent).

Theorem 16 (Inverse)**.** If a matrix A has both a left inverse B and a right inverse C, then $B = C$ and A is square and row equivalent to I. Then A is *invertible* and has *inverse* A^{-1}.

Proof. $B = BAC = C$ □

Definition 11. The *transpose* of a matrix is $A^\top$ where $(A^\top)_{ij} = A_{ji}$.

Remark 6. The transpose satisfies $(AB)^\top = B^\top A^\top$. If A is invertible, then $(A^\top)^{-1} = (A^{-1})^\top$ since $(AA^{-1})^\top = I = (A^{-1})^\top A^\top$.

Definition 12. The *coordinates* of x in basis $\mathcal{B}$ are $x_{\mathcal{B}}$ s.t. $x = \mathcal{B}x_{\mathcal{B}}$.

Theorem 17. $\mathcal{B}$ is an isomorphic map from coordinates to vectors, while $\mathcal{B}^{-1}$ translates vectors to coordinates. The *change-of-basis* formula $\underset{\mathcal{C}\leftarrow\mathcal{B}}{L} = \mathcal{C}^{-1}\mathcal{B}$ converts coordinates in $\mathcal{B}$ to coordinates in $\mathcal{C}$, and $\mathcal{T}^{-1}A\mathcal{S}$ converts a matrix A on standard coordinates to $\mathcal{S} \to \mathcal{T}$.

Proof. The new domain is $\mathcal{S}^{-1}\mathbf{x}$ and range is $\mathcal{T}^{-1}A\mathbf{x} = (\mathcal{T}^{-1}A\mathbf{x}\mathcal{S})(\mathcal{S}^{-1}\mathbf{x})$. □

4. Size

Theorem 18. The *determinant* $|A|$ represents the volume scaling factor of a transformation or the signed volume of the columns of its matrix.
Axiomatically, the determinant is the unique alternate multilinear function s.t. $|I| = 1$:

$$|A| = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$$

where the sum is taken over all permutations of 1 through n.

- A matrix with a zero row has a determinant of 0
- A singular matrix has a determinant of 0, since there is a linear dependence
- $|A| = |A^\top|$

Theorem 19 (Laplacian expansion by minors)**.**

$$|A| = \sum_{i=1}^{k} C_{ij}$$

Cofactor $C_{ij} = (-1)^{i+j}M_{ij}$ and minor M_{ij} is the determinant of A less row i and column j.

Theorem 20. Elementary matrices have the following determinants:

- 1 for row replacement
- -1 for interchange (alternate)
- k for scaling by factor k

For invertible A, $|A| = |E_1E_2\cdots E_nI| \neq 0$, so

- The determinant is multiplicative: $|AB| = |A||B|$
- $AA^{-1} = I \Rightarrow |A||A^{-1}| = 1$
- The determinant is the product of the pivots if the matrix is invertible
- The determinant of a triangular matrix is the product of the diagonal entries

Theorem 21 (Cramer's Rule)**.** The solutions to $A\mathbf{x} = \mathbf{b}$ are

$$x_i = \frac{|A_i(\mathbf{b})|}{|A|}$$

where $A_i(\mathbf{b})$ is the matrix which replaces the ith column of A with $\mathbf{b}$.

Proof. $A\mathbf{x} = \mathbf{b} \Rightarrow AI_i(x) = A_i(\mathbf{b}) \Rightarrow |A|x_i = |A_i(\mathbf{b})|$ □

Theorem 22. Applying Cramer's rule against $A\mathbf{x} = \mathbf{e}_i$ shows that

$$A^{-1} = \frac{\operatorname{adj}(A)}{|A|}$$

where the adjugate matrix is the transpose of the cofactor matrix.

Definition 13. An eigenvalue λ and corresponding eigenvector $\mathbf{v}$ satisfies

$$A\mathbf{v} = \lambda\mathbf{v} \qquad (\mathbf{v} \neq 0)$$

Any solution must satisfy $(A - \lambda I)\mathbf{v} = \mathbf{0}$, so $A - \lambda I$ must be singular, and eigenvalues are roots of the characteristic polynomial $\chi = |A - \lambda I|$. In particular the product of eigenvalues is $|A|$ and the sum is $\text{Tr}(A)$.

Corollary 5. A matrix is singular iff it has a zero eigenvalue, since then $A - \lambda I = A$ has a nontrivial nullspace.

Theorem 23. The eigenvalues of a triangular matrix are the diagonal entries, since $A - \lambda I$ is still triangular and the determinant is just the product of the main diagonal.

Theorem 24. The algebraic multiplicity of an eigenvalue λ is its multiplicity as a root of the charpoly. The geometric multiplicity is the dimension of the eigenspace $A - \lambda I$, which is at least 1 (since $A - \lambda I$ is singular) and at most the algebraic multiplicity.

Proof. Suppose the geometric multiplicity of λ is k, and let $U = u_1, \cdots, u_k$ be the basis for the corresponding subspace. Extend this basis to the entire space, and let B be the change-of-basis matrix. Then the charpoly of A is the charpoly of $Q^{-1}AQ$, which is a multiple of $(t - \lambda)^k$ since the first k elements are λI_k. □

Proof. Let the algebraic multiplicity be k. Then $\text{rref}(A - \lambda I)$ is upper-triangular, so it has 0 eigenvalue at most k times, so it has at most k zero rows, so the dimension of the eigenspace is at most k. □

Theorem 25. Eigenvectors corresponding to distinct eigenvalues are independent.

Proof. Suppose there is a dependence $\mathbf{v} = c_1\mathbf{v_1} + \cdots + c_k\mathbf{v_k}$ for nonzero c_i. Multiplying by A, $\lambda\mathbf{v} = c_1\lambda_1\mathbf{v_1} + \cdots + c_k\lambda_k\mathbf{v_k}$. Multiplying by λ, $\lambda\mathbf{v} = c_1\lambda\mathbf{v_1} + \cdots + c_k\lambda\mathbf{v_k}$. Thus $0 = c_1(\lambda - \lambda_1)\mathbf{v_1} + \cdots + c_k(\lambda - \lambda_k)\mathbf{v_k}$, a contradiction. □

Definition 14. Square matrices A and B are *similar* if invertible P exists s.t. $B = P^{-1}AP$.

- $\chi = |B - \lambda I| = |P^{-1}AP - P^{-1}\lambda IP| = |P^{-1}(A - \lambda I)P| = |A - \lambda I|$, so similar matrices have the same charpoly, eigenvalues, determinant, and trace. Of course, the eigenvectors are different.
- A^{-1} exists iff B^{-1} exists, and $B^{-1} = P^{-1}A^{-1}P$. This is equivalent to having the same charpoly and being diagonalizable.
- If $\mathbf{v}$ is an eigenvector of A, then $P^{-1}\mathbf{v}$ is an eigenvector of B, since $B(P^{-1}\mathbf{v}) = (P^{-1}AP)P^{-1}\mathbf{v} = P^{-1}A\mathbf{v} = \lambda(P^{-1}\mathbf{v})$

Definition 15. A matrix is *diagonalizable* if it is similar to a diagonal matrix D, that is $A = PDP^{-1}$. Equivalently, the eigenvectors span $\mathbb{R}^n$, since then the eigenbasis P and their corresponding eigenvectors D satisfy $AP = PD \Rightarrow A = PDP^{-1}$. A matrix is diagonalizable over reals if all eigenvalues are real and the dimension of each eigenspace equals the multiplicity of its eigenvector.

5. Orthogonality

Definition 16. The *inner product* or *dot product* $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^\top\mathbf{v}$ is symmetric, bilinear, and positive definite, that is, $\mathbf{u} \cdot \mathbf{u} \geq 0$, with equality when $\mathbf{u} = \mathbf{0}$. The *length* or *norm* of a vector is $|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. The distance between $\mathbf{u}$ and $\mathbf{v}$ is $|\mathbf{u} - \mathbf{v}|$. $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* iff

$\mathbf{u} \cdot \mathbf{v} = 0$, or equivalently if $|\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2$. The *orthogonal complement* of a space V is the set of vectors orthogonal to every vector in V.

Theorem 26 (Cauchy-Schwarz)**.** $\mathbf{x} \cdot \mathbf{y} \leq |x||y|$, with equality when $\mathbf{x} = k\mathbf{y}$.

Theorem 27. An orthogonal set of nonzero vectors is linearly independent.

Proof. Suppose that $c_1\mathbf{v_1} + \cdots + c_n\mathbf{v_n} = 0$. Then $0 = 0 \cdot \mathbf{v_1} = (c_1\mathbf{v_1} + \cdots + c_n\mathbf{v_n}) \cdot \mathbf{v_1} = c_1(\mathbf{v_1} \cdot \mathbf{v_1})$, so $c_1 = 0$ since $\mathbf{v_1} \neq 0$. Similarly, $c_i = 0$. □

Theorem 28. The coordinates of v in an orthogonal basis U are

$$c_i = \frac{\mathbf{v} \cdot \mathbf{u_i}}{\mathbf{u_i} \cdot \mathbf{u_i}}$$

The *orthogonal projection* of $\mathbf{v}$ onto $\mathbf{u}$ $\text{proj}_{\mathbf{u}}\mathbf{v}$ satisfies $\text{proj}_{\mathbf{u}}\mathbf{v} = k\mathbf{u}$ and $(\mathbf{v} - \text{proj}_{\mathbf{u}}\mathbf{v}) \cdot \mathbf{u} = 0$

$$\text{proj}_{\mathbf{u}}\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}$$

The *orthogonal projection* of $\mathbf{v}$ onto a subspace W with orthogonal basis U is

$$\text{proj}_W\mathbf{v} = \sum \frac{\mathbf{v} \cdot \mathbf{u_i}}{\mathbf{u_i} \cdot \mathbf{u_i}}$$

The *Gram-Schmidt process* orthonormalizes a basis.

Definition 17. The *least squares solution* $\hat{\mathbf{x}}$ to $A\mathbf{x} = \mathbf{b}$ minimizes $|A\hat{\mathbf{x}} - \mathbf{b}|$. It satisfies the *normal equation* $A^\top A\hat{\mathbf{x}} = A^\top\mathbf{b}$ since $|\mathbf{v} - \text{proj}_W\mathbf{v}| < |\mathbf{v} - \mathbf{y}|$ for all other $\mathbf{y}$ in W.

Definition 18. An *orthogonal matrix* is a square matrix U with orthonormal columns or orthonomal rows, or equivalently $UU^T = U^TU = I$ or $U^T = U^{-1}$. They represent rigid rotations and reflections which preserve lengths and angles:

- $|U\mathbf{x}| = |\mathbf{x}|$
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$

For an orthogonal basis,

$$\text{proj}_U\mathbf{v} = \sum(\mathbf{v} \cdot \mathbf{u_i})\mathbf{u_i} = UU^\top\mathbf{v}$$

The eigenvalues of an orthogonal matrix have magnitude 1, so the determinant is ± 1. The product of two orthogonal matrices is also orthogonal.

Definition 19. Matrix A is *symmetric* if $A^\top = A$.

Theorem 29 (Spectral Theorem)**.** A is orthogonally diagonalizable iff A is symmetric.

Proof. $A^\top = (PDP^\top)^\top = PDP^\top = A$.
If $\mathbf{v}_1$ and $\mathbf{v}_2$ correspond to distinct eigenvalues, then $\lambda_1\mathbf{v}_1 \cdot \mathbf{v}_2 = (\lambda_1\mathbf{v}_1)^\top\mathbf{v}_2 = (A\mathbf{v}_1)^\top\mathbf{v}_2 = (v_1^\top A^\top)\mathbf{v}_2 = v_1^\top A\mathbf{v}_2 = v_1^\top\lambda_2\mathbf{v}_2 = \lambda_2\mathbf{v}_1 \cdot \mathbf{v}_2$. Thus $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, so $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. □

Definition 20. The *spectral decomposition* of A is

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^\top + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^\top$$

6. Differential Equations

The **first-order homogenous ODE** $y' + ay = 0$ with the initial value $y(0)$ has the solution

$$y_h = y(0)e^{-at}$$

The **first-order inhomogenous ODE** $y' + ay = f(t)$ has the solution $y = y_p + y_h$.
Method of undetermined coefficients. If $f(t) = Ct^m e^{rt}$, then

$$y_p = (A_m t^m + \cdots + A_0)e^{rt}$$

and the coefficients are found by plugging it into the ODE.
If $r = -a$, multiply by t: $y_p = t(A_m t^m + \cdots + A_0)e^{rt}$.
If there is a sine or cosine term, then Euler's formula shows that
$y_p = (A_m t^m + \cdots + A_0)e^{\alpha t}\cos\beta t + (B_m t^m + \cdots + B_0)e^{\alpha t}\sin\beta t$.

Superposition principle. If y_1 is a solution to $y' + ay = f_1(t)$ and y_2 is a solution to $y' + ay = f_2(t)$, then $k_1y_1 + k_2y_2$ is a solution to $y' + ay = k_1f_1(t) + k_2f_2(t)$. The y_p for $f(x)$ containing sums is the sum of the y_p for each term.

The **homogenous linear system** $\mathbf{x}' = A\mathbf{x}$ has the solution space $\mathrm{Span}\lambda_i\mathbf{v}_i$, which may be written as a *fundamental matrix* $X = [\lambda_i\mathbf{v}_i]$.
The solution may also be written as the *matrix exponential* $e^{At} = XX^{-1}(0)$ where $e^{A(0)} = I$.
The two-dimensional case may be visualized with a *phase portrait*, which is a vector field with value (x', y') at each (x, y). The behavior along eigenvectors and axes is important.

The **inhomogenous linear system** $\mathbf{x}' = A\mathbf{x} + \mathbf{b}$ has a vector y_p such as $\mathbf{a}e^{rt}$.
Or use **variation of parameters**: $\mathbf{x} = e^{A(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^{t} e^{A(t-t_0)}\mathbf{f}(s)\,\mathrm{d}s$.

The **higher-order ODE** $y^{(n)} + \cdots + a_1y' + a_0y = f(t)$ may be written in *normal form* as a first-order system using the substitution $x_0 = y$, $x_1 = y'$, …, $x_{n-1} = y^{(n-1)}$, so

$$\mathbf{x}' = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix}\mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(x) \end{bmatrix}$$

The **Fourier series** for $f(x)$ on the interval $[-L, L]$ is

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi}{L}x + b_n \sin\frac{n\pi}{L}x, \text{where}$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi}{L}x\,\mathrm{d}x \qquad b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi}{L}x\,\mathrm{d}x$$

A cosine or sine series may be written on the *half-range* $[0, L]$ assuming that the function is even or odd, respectively.

To solve a **PDE** such as the heat or wave equation, expand the solution u as a Fourier series in the eigenbasis that satisfies the boundary conditions. Add a $U_1 + \frac{U_2-U_1}{2}x$ term if the boundary is inhomogenous. Then write out the main equation, expressing the inhomogenous term in the eigenbasis. Collecting terms yields an ODE system. Finally, ensure that the solution satisfies the initial conditions.