Eigenvalues and eigenvectors:

- **Eigenvalues** for a matrix A are the solutions to the equation $\det(A - \lambda I) = 0$.
  - Σ eigenvalues = trace of the matrix.
    - The trace of a square matrix is the sum of all its diagonal elements
  - Π eigenvalues = determinant of matrix
  - The matrix is invertible if Π eigenvalues $\neq 0$
- **Eigenvectors** are **linearly independent vectors** that represent the nullspace of $A - \lambda I$.
  - To find all eigenvectors for a matrix:
    - Find all eigenvalues of the matrix A by solving $\det(A - \lambda I) = 0$.
    - For each one of them (consider repetitions), find the corresponding eigenvector by solving the equation $(A - \lambda I)\, x = \mathbf{0}$
- If $\mathbf{Ax = \lambda x}$ then λ is an **eigenvalue** and x is an **eigenvector** of the matrix A.
  - $\mathbf{Ax = \lambda x, Bx = \alpha x}$ does not necessarily imply that $\alpha + \lambda$ is an eigenvalue of A+B
    - It's true of A and B share the same eigenvectors
    - Same with multiplication.
  - If the matrix has a **complex eigenvalue**, then its conjugate is also an eigenvalue of the same matrix.
  - **Symmetric matrices** have real eigenvalues and perpendicular eigenvectors.
    - Because the eigenvectors are perpendicular, then the matrix S (details are next) is orthogonal matrix and is denoted by Q
      - It's known that $Q^{-1} = Q^T$, this will help in diagonalization process.
      - Q is the same as S, but its column vectors have to be **normalized**.
  - **Antisymmetric** matrices ($A^T = -A$) have all eigenvalues imaginary (in form ci).
  - In **triangular** matrices, eigenvalues are the elements on the main diagonal.
  - Eigenvalues for A^k = eigenvalues for A, but raised to the power k.
  - Eigenvalues of A^T are the same as eigenvalues of A.

Matrix diagonalization: $A = S\Lambda S^{-1}$

- The process of converting A into a diagonal matrix which shares the same characteristics with A.
- This iterative algorithm was discussed in the tutorial, but not in MIT lecture notes.
  - https://en.wikipedia.org/wiki/Jacobi_eigenvalue_algorithm
- The straightforward algorithm is to find all eigenvalues and eigenvectors of A, then:
  - Construct the matrix S where S is the matrix with all linearly independent eigenvectors of A as columns. $S = [\mathbf{x_1} \quad \mathbf{x_2} \quad \ldots \quad \mathbf{x_n}]$
  - Λ is the diagonal matrix with all distinct eigenvalues at the main diagonal
  - Then we have: $S^{-1}AS = \Lambda$, $A^k = S\Lambda^k S^{-1}$
- The matrix A of dimension n*n is diagonalizable iff it has n linearly independent eigenvectors.
  - The **algebraic** multiplicity of an eigenvalue e is the power to which $(\lambda - e)$ divides the characteristic polynomial.
  - The **geometric** multiplicity of an eigenvalue e is the number of linearly independent eigenvectors associated with it.
  - Distinct eigenvalues imply linearly independent eigenvectors, the converse is not always true.
  - A is diagonalizable iff for every eigenvalue that has multiplicity k, we have k linearly independent corresponding eigenvectors.

Applications of eigenvalues and vectors:

- The system of n differential equation in the form:

$$\begin{cases} \frac{du_1}{dt} = \sum_{i=1}^{n} \mathrm{a}_{1\mathrm{i}} u_i \\ \quad \cdots \\ \frac{du_n}{dt} = \sum_{i=1}^{n} \mathrm{a}_{\mathrm{ni}} u_i \end{cases}$$ is equivalent to $\frac{d\boldsymbol{u}}{dt} = A\boldsymbol{u}$ where A = $\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$, **u** = $\begin{bmatrix} u_1 \\ .. \\ u_n \end{bmatrix}$

- The general solution for this system will be:

$$\boldsymbol{u}(t) = e^{At}u(0) = Se^{\Lambda t}S^{-1}u(0), e^{\Lambda t} = \begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{\lambda_n t} \end{pmatrix}$$

- Real parts of eigenvalues of A determine the stability of the solution,
  - If Re(λ) > 0 for at least 1 λ then the solution is **not stable** and has no steady state.
  - Else if Re(λ) < 0 for all λ's then solution is stable and u(t) → 0 when t → ∞
  - Else if at least one λ = 0 (all others are negative) then the system has a steady state and u(∞) approaches some finite value.
- We don't need to calculate eigenvalues for the matrix to do the stability test, alternatively we use matrix pivots:
  - Signs of matrix pivots = signs of its eigenvalues.

Positive definite matrices:

- A positive definite matrix (A) is the matrix that satisfy any of these properties:
  - $x^TAx > 0$ for any nonzero real vector x.
    - if $x^TAx = 0$ then the matrix is positive semi-definite.
  - All the eigenvalues of M are strictly positive.
  - All the upper-left submatrices have positive determinants (Sylvester criterion).
  - All the pivots (without row exchanges) are strictly positive.
    - The k-th pivot $d_k = \det(A_k)/\det(A_{k-1})$ "upper left submatrices", $\det(A_0) = 1$.
    - The pivots are the nonzero elements of D in the LDL^T decomposition.
    - The pivots are the elements that are used in Gaussian elimination
    - The pivots are the nonzero leading elements in REF.
  - It can be decomposed to $A = R^TR$, R has independent column vectors.
- **Property:** if A and B are both symmetric and positive definite, then so is A+B.
- **Before discussing applications, we need to know that some polynomials have a corresponding symmetric matrix that is used to know some of their characteristics.**
  - Let **x** = $[x_1 \;\; x_2 \; \ldots \; x_n]$ be a vector of variables.
- **Examples:**
  - For $\mathrm{P}(\mathbf{x}) = \mathrm{a}\mathrm{x}_1^2 + \mathrm{b}\mathrm{x}_2^2 + 2\mathrm{c}\,\mathrm{x}_1\mathrm{x}_2$
    - The corresponding matrix $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$.
  - For $\mathrm{P}(\mathbf{x}) = \mathrm{a}\mathrm{x}_1^2 + \mathrm{b}\mathrm{x}_2^2 + \mathrm{c}\mathrm{x}_3^2 + 2\mathrm{h}\,\mathrm{x}_1\mathrm{x}_2 + 2\mathrm{g}\,\mathrm{x}_1\mathrm{x}_3 + 2\mathrm{f}\,\mathrm{x}_2\mathrm{x}_3$
    - The corresponding matrix $A = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$.

Applications of symmetric positive definite matrices.

- Solving $P(\mathbf{x}) > 0$ is the same as determining if the corresponding matrix A is positive definite, we can use eigenvalues test or Sylvester criterion.
  - Same for solving $P(\mathbf{x}) < 0$ and negative definite
- For finding minima:
  - Let $P(\mathbf{x})$ be a polynomial that has a corresponding positive definite matrix, then:
    - $A\mathbf{x} = b \Leftrightarrow P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^TA\mathbf{x} - \mathbf{x}^Tb$.
    - $A\mathbf{x} = b$ is the SLE that results from setting all the partial derivatives to 0, A is the corresponding matrix for the polynomial P.
    - We have $P_{min}(\mathbf{x}) = -\frac{1}{2}b^TA^{-1}b$, the minimum value of the polynomial.
- To check if a point $P(x, y)$ is a minima or maxima or saddle point for a multivariable function $f(\mathbf{x})$
  - Construct the Hessian matrix of second partial derivatives.
    - If H is positive definite → P is a local min.
    - If H is negative definite → P is a local max.
    - If H is Indefinite → P is a saddle point.
    - If H is ± semi-definite → test is inconclusive.

$$\mathbf{H} = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

Theorem 1.1 (Sylvester's criterion). *Let A be an $n \times n$ symmetric matrix. Then:*

- $A \succ 0$ *if and only if* $\Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_n > 0$.
- $A \prec 0$ *if and only if* $(-1)^1\Delta_1 > 0, (-1)^2\Delta_2 > 0, \ldots, (-1)^n\Delta_n > 0$.
- *A is indefinite if the first Δ_k that breaks both patterns is the wrong sign.*
- *Sylvester's criterion is inconclusive (A can be positive or negative semidefinite, or indefinite) if the first Δ_k that breaks both patterns is 0.*

- To find the axes of rotation of a quadratic curve.
  - Represent the equation of the curve as $\mathbf{x}^TA\mathbf{x} = c$
  - Decompose A into $Q\Lambda Q^T$.
  - Now we have the original equation as $(\mathbf{x}^TQ)\Lambda(Q^T\mathbf{x}) = c \rightarrow X^T\Lambda X = c$
  - Construct the new polynomial from that system in the form of $\lambda_1u^2 + \lambda_2v^2 = c$
  - Axes of rotation are $u = 0$, $v = 0$.

Similar matrices:

- A and B are similar if $B = M^{-1}AM$, for some non-singular matrix M.
- The matrices that have the same eigenvalues are all similar to each other and similar to Λ.
  - Because all of them can be decomposed into $A = S\Lambda S^{-1}$
  - They don't necessarily have the same eigenvectors.

Singular value decomposition: $A = U\Sigma V^T$

- U: eigenvectors matrix (S) of AA^T, V: eigenvectors matrix (S) of A^TA
- Both U and V are orthogonal matrices ⇒ $UU^T = I = VV^T$
  - All v_i and u_i (eigenvectors) should be normalized.
- Σ is a diagonal scale matrix = the sqrt of eigenvectors of A^TA or AA^T
- $\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_r \end{bmatrix} = \sqrt{\Lambda}$, Λ = eigenvalues(A^TA) = eigenvalues(AA^T)
  - The last equality doesn't hold for non-square matrices.

- **Informal description of calculating SVD for non-square matrices:**
  - A^TA and AA^T will still be square matrices, but they will not have the same dimensions. We can still find U (dimension a*b) and V (dimension c*d).
  - In order for the equality $\mathbf{A = U\Sigma V^T}$ to hold, Σ should have dimension b*c.
    - If needed, add more zero rows/columns to satisfy that requirement.

SVD and the fundamental spaces:

- **From the fact that $Av_i = \sigma_i u_i$ we conclude (r = Rank(A), A is n*n matrix)**
  - $v_1, v_2, \ldots, v_r$ is an orthonormal bases for the row space of A.
  - $u_1, u_2, \ldots, u_r$ is an orthonormal bases for the column space of A.
  - $v_{r+1}, \ldots, v_n$ is an orthonormal bases for the null space of A.
  - $u_{r+1}, \ldots, u_n$ is an orthonormal bases for the left null space of A.
    - if r = n, the null space and the left null space contain only the zero vector.

Left and right inverse

- For square non-singular matrices, the left and right inverse are the same.
- For rectangular matrices, the left and right inverse cannot be the same.
  - Consider the rectangular matrix A with dimension n*m.
    - $A_{\text{left}}^{-1} = (A^TA)^{-1}A^T$ is the matrix that satisfies the property $A_{\text{left}}^{-1}\,A = I_m$.
    - $A_{\text{right}}^{-1} = A^T(AA^T)^{-1}$ is the matrix that satisfies the property $AA_{\text{right}}^{-1} = I_n$.
      - These matrices are not necessarily unique.

Pseudoinverse

- Suppose we have several distinct vectors $x_1, x_2, \ldots, x_n$ and we multiplied each of them from the left by a non-zero matrix A to get $Ax_1, Ax_2, \ldots, Ax_n$
- This operation is invertible, by multiplying the resulting vectors from the left by the pseudoinverse of A.
- The pseudoinverse for a matrix A is the matrix A^+ s.t $\mathbf{x} = A^+A\mathbf{x}$ $\forall x \in$ Rowspace(A).
- AA_{left}^{-1} and $A_{\text{right}}^{-1}A$ are both possible pseudoinverses for A.
- If $A = U\Sigma V^T$ then $A^+ = V\Sigma^+U^T$, Σ^+ is the same as Σ with every σ_i inverted to be $1/\sigma_i$.
  - If $\sigma_i = 0$, then we keep it 0 while calculating Σ^+