Lecture 1

1. Linear algebra aims to solve linear equation systems. For example,

$$\begin{cases} 2x - y = 0 \\ -x + 2y = 3 \end{cases}$$

can be written as

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Therefore, matrix times column vector Ax is a linear combination of columns of A.
Similarly, row vector times matrix $x^T A$ is a linear combination of rows of A.

Lecture 2 - 4

1. We solve linear equation systems by row elimination to convert A to an upper triangle matrix U.
   Row elimination (row operation) is equivalent to left multiply by a n elimination matrix E.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 2 & 5 & 8 \end{bmatrix} \rightarrow E_{21}A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 2 & 5 & 8 \end{bmatrix} \rightarrow E_{31}E_{21}A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \rightarrow E_{32}E_{31}E_{21}A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$where\ E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

   Notice that:

   (a) E_{ij} is lower triangle with only 1 non-zero element at (i,j) position, which subtracts some multiples of row j from row I, and makes (i,j) entry of A zero.
   (b) To calculate E^{-1}, just flip the off-diagonal numbers. E^{-1} is also an elimination matrix.
   (c) Row elimination is essentially LU decomposition
2. Sometimes row exchange is required in the process of LU decomposition, then $A = LU$ becomes $PA = LU, wher\ P\ is\ a\ permutation\ matrix$.
   Permutation matrix P_{ij} is the identity matrix with $row\ i\ and\ j$ reversed. $P^{-1} = P^T$.
3. $A_{n\times n}$ is invertible is equivalent to:
   (a) A has n pivots after elimination (row exchange allowed)
4. Diagonally dominant matrices are invertible. Diagonally dominant matrices are defined as $|a_{ii}| > \sum_{j\neq i}|a_{ij}|\ for\ all\ i$
5. Inverse of upper triangle is upper triangle; Inverse of lower triangle is lower triangle.

Lecture 5

1. Subspace: a set with components closed under addition and scalar multiplication.
2. Possible subspaces in R^2: $(1)R^2$; $(2)any\ line\ through\ the\ origin$; (3) $\{(0,0)\}$.
3. Column space: all linear combination of columns of A, denoted as $C(A)$.
4. Null space: $\{x|Ax = 0\}, denoted\ as\ N(A)$.

Lecture 6

1. $Ax = b$ can be solvable when $b \in C(A)$.

2. The solution of $Ax = b$ is $a\ particular\ solution + N(A), i.e., a\ shifted\ subspace$.

Lecture 7

1. Reduced row echelon form: make all pivot elements 1.
2. $A \sim R = \begin{bmatrix} I_{r\times r} & F_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$
   $then\ Ax = 0 \Leftrightarrow Rx = 0 \Rightarrow N(A) = C\left(\begin{bmatrix} -F_{r\times(n-r)} \\ I_{(n-r)\times(n-r)} \end{bmatrix}\right) = C\left(\begin{bmatrix} -F \\ I \end{bmatrix}\right)$

Lecture 8-9

1. For a $m \times n$ matrix A, we have $r \le m, r \le n$.
   - Full column rank: r=n. No free variables. All columns intendent.
     $N(A) = \{\vec{0}\}$.
     $Ax = b$ has either no solution or unique solution.
   - Full row rank: $r = m.\ \ A\ has\ r\ pivots, (n - r)\ free\ variables$.
     $Ax = b$ always solvable, either 1 or ∞ solutions.
   - Full rank: $r = m = n. A$ is invertible, $R = I$.
     $Ax = b$ always has unique solution $x = A^{-1}b$
   - $r < m, r < m.$
     $Ax = b\ has\ either\ 0\ or\ \infty\ solutions.$
2. $r = m = n, 1\ solution$
   $r = m < n, \infty\ solutions$
   $r = n < m, 0\ or\ 1\ solution$
   $r < m, r < n, 0\ or\ \infty\ solutions$
3. $rank(A) = \dim\big(C(A)\big) = \dim\ (rowspace(A))$

Lecture 10

1. Four basic subspaces

	Which space is it in?	$dimension$	$find\ one\ basis$
$row\ space$	R^n	r	First r rows of rref
$null\ space$ $N(A)$	R^n	$n - r$	Special solutions
$column\ space$ $C(A)$	R^m	r	Pivot columns
$left\ null\ space$ $N(A^T)$	R^m	$m - r$	Last (m-r) rows of E, where EA=R $[A\ \ \ I] \sim [R\ \ \ E]$

2. Orthogonal subspaces: $v^T w = 0\ for\ all\ v \in V\ and\ all\ w \in W$
3. Orthogonal complement of subspace V contains every vector perpendicular to V.
4. $row\ space\ \perp null\ space$
   $column\ space \perp left\ null\ space$
   $also\ orthogonal\ complements$
5. Every vector can be split into a row space component and a null space component.

The basis for row space + basis for null space span R^n

Lecture 15

1. Project a vector $\vec{b}$ onto $\vec{a}$
$$\begin{cases} p = xa \\ a^T(b-p) = 0 \end{cases} \Longrightarrow a^R b = x a^T a \Longrightarrow x = \frac{a^T b}{a^T a}, p = \frac{a a^T b}{a^T a}$$
$Let\ p = Pb, we\ get\ the\ projection\ matrix\ P = \frac{aa^T}{a^T a}$
2. Projection & least square: sometimes $Ax = b$ may have no solution, so solve $A\hat{x} = p$ $instead, where\ p\ is\ projection\ of\ b\ onto\ C(A).$
3. Projection b onto a plane spanned by $a_1, a_2. Let\ A = [a_1\ a_2], projection\ p = A\hat{x}$
$$a_1^T(b - A\hat{x}) = 0, a_2^T(b - A\hat{x}) = 0$$
$$\Leftrightarrow \begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix}(b - A\hat{x}) = 0 \Leftrightarrow A^T(b - A\hat{x}) = 0$$
   (1) $e = b - A\hat{x}\ is\ in\ N(A^T) \Leftrightarrow e \perp C(A)$
   (2) $\hat{x} = (A^T A)^{-1} A^T b$
   (3) $p = A\hat{x} = A(A^T A)^{-1} A^T b$
   (4) $projection\ matrix\ P = A(A^T A)^{-1} A^T$
   (5) $P\ has\ properties: P^T = P, P^2 = P$

Lecture 16

1. Projection matrix $P = A(A^T A)^{-1} A^T$
$If\ b \in C(A), then\ b = Ax, thus\ p = Pb = A(A^T A)^{-1} A^T Ax = Ax = b \Rightarrow projection\ is\ itself$
$If\ b \perp C(A), then\ A^T b = 0, thus\ p = Pb = A(A^T A)^{-1} A^T b = 0 \Rightarrow peojection\ is\ 0$
2. $b = projection + error = Pb + (I - P)b,$
$where\ P\ is\ projection\ matrix\ on\ to\ C(A), \quad I - P\ is\ projection\ matrix\ onto\ N(A^T)$
3. $I - P\ is\ also\ a\ projection\ matrix: (I - P)^T = I - P, (I - P)^2 = I - P$
4. $For\ any\ matrix\ A, A\ and\ A^T A\ has\ the\ same\ null\ space.$
$Proof: For\ one\ side, Ax = 0 \Rightarrow A^T Ax = 0$
$For\ the\ other\ side, A^T Ax = 0 \Rightarrow x^T A^T Ax = 0 \Rightarrow (Ax)^T(Ax) = 0 \Rightarrow Ax = 0$
$Thus\ \forall A, Ax = 0 \Leftrightarrow A^T Ax = 0.$
5. $If\ A\ has\ indepedent\ columns, then\ A^T A\ is\ invertible.$
$Proof: \because A\ has\ indepedent\ columns, \therefore N(A)\ only\ contains\ 0.$
$From\ the\ lemma\ in\ 4, N(A^T A)\ only\ contains\ 0\ as\ well. Thus\ A^T A\ is\ invertible.$

Lecture 17 Orthogonal Matrix

1. Vectors $q_1, q_2, \ldots, q_n$ are orthonormal if $q_i^T q_j = \begin{cases} 0, & if\ i \neq j \\ 1, & if\ i = j \end{cases}.$
$Let\ Q = [q_1\ \ q_2\ \ \ldots\ \ q_n]$, then $Q^T Q = I_n$
$If\ Q\ is\ also\ square, then\ QQ^T = I, and\ Q\ is\ called\ an\ orthogonal\ matrix.$
2. Examples of orthogonal matrix:
   (1) Permutation matrix: $Q = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

(2) Rotation matrix: $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

(3) Reflection matrix: $Q = I - 2uu^T, where\ u\ is\ any\ unit\ vector\ \ (uu^T\ is\ projection\ matrix)$

3. Why orthogonal matrix better than ordinary matrix?

   (1) Projection matrix $P = Q(Q^TQ)^{-1}Q^T = QQ^T$

   (2) OLS: $Q^TQ\hat{x} = Q^Tb \Leftrightarrow \hat{x} = Q^Tb$

4. Gram-Schmidt: convert a set of independent vectors $x_1, x_2, \dots, x_n$ to a set of orthonormal vectors $q_1, q_2, \dots, q_n$

5. $Gram\ Schmidt: A = QR,\ where\ Q\ is\ an\ orthogonal\ matrix, R\ is\ an\ upper\ triangle\ square\ matrix.$

   (remember $A = LU$ is row elimination)

6. Any matrix with independent columns can do QR decomposition.

Lecture 18 Determinant

We use 3 properties to define determinant:

1. $\det(I) = 1$
2. $Exchange\ 2\ rows\ reverse\ the\ sign$
3. $Linearity\ for\ each\ row$

   (1) $\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t\begin{vmatrix} a & b \\ c & d \end{vmatrix}$

   (2) $\begin{vmatrix} a + a' & b + b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$

Using the above 3 properties, we can prove more properties.

4. $If\ A\ has\ 2\ equal\ rows, then\ |A| = 0. (From\ property\ 2)$
5. $Subtract\ i \times row\ j\ from\ row\ k \to determinant\ not\ change\ (From\ property\ 3 + 4)$
6. $If\ A\ has\ a\ row\ of\ 0 \to |A| = 0\ (From\ property\ 4\ and\ 5)$
7. $U = \begin{bmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \to \det(U) = d_1 d_2 \dots d_n\ (From\ property\ 1 + 3.1 + 5)$
8. $\det(A) = 0 \Leftrightarrow A\ is\ singular$

   Proof: from property 2 and 5, if we do LU decomposition $PA = LU$, then $\det(A) = \pm\det(U)$. Therefore, $\det(A) = 0 \Leftrightarrow \det(U) = 0 \Leftrightarrow (property\ 7) U\ has\ a\ row\ of\ 0 \Leftrightarrow A\ singular.$
9. $\det(AB) = \det(A)\det(B)$. Proof is a little tricky. Convert A=LU (if row exchange required then $A = P^TLU$, and remaining is same), then$\det(AB) = \det(LUB) = \det(UB) = d_1 d_2 \dots d_n \det(U'B) = d_1 d_2 \dots d_n \det(B) = \det(A)\det(B)$. Here we use the fact that $\det(LA) = \det(A)\ if\ L\ is\ a\ lower\ triangle\ with\ diagnal = 1 (from\ preperty\ 5)$.
10. $\det(A^T) = \det(A)$.

    $Proof: \det(A^T) = \det(U^TL^T) = \det(U^T)\det(L^T) = \det(L)\det(U) = \det(LU) = \det(A)$.

Lecture 20

1. $A^{-1} = \frac{C^T}{\det(A)}, where\ C^T\ is\ the\ matrix\ of\ cofactors.$

$Proof: check\ AC^T = \det(A)\, I.$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix}$$

$diagonal: a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} = \det(A)$

$off\ diagonal: a_{21}C_{11} + a_{22}C_{12} + \cdots + a_{2n}C_{1n} = \det\left(\begin{bmatrix} a_{21} & \cdots & a_{2n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}\right) = 0$

2. $Cramer's\ rule: If\ \det(A) \neq 0, then\ Ax = b\ can\ be\ solved\ by\ determinants:$
   $x_1 = \frac{\det(B_1)}{\det(A)}, x_2 = \frac{\det(B_2)}{\det(A)}, \ldots, x_n = \frac{\det(B_n)}{\det(A)}$
   $where\ B_j\ has\ the\ jth\ column\ of\ A\ relaced\ by\ b.$
   $Proof: A\begin{bmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{bmatrix} \Rightarrow \det(A) \cdot x_1 = \det(B_1) \Rightarrow x_1 = \frac{\det(B_1)}{\det(A)}$
   Note: Cramer's rule is computationally expensive and thus not used in practice.

3. Cross product of $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ is
   $$u \times v = \begin{vmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$
   a. $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin\theta|, direction\ by\ right\ hand\ rule$
   b. $u \times v \perp u, u \times v \perp v$
4. Determinant and area/volume
   a. In 2-D plane, if a parallelogram has two edges $x = (x_1, x_2)\ and\ y = (y_1, y_2)$, then the area is abs of $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$
   b. In 3-D space, if a parallelepiped has three edges u, v, w, then the volume is abs of $\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$
5. Triple product: $(u \times v) \cdot w = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$

Lecture 21 Eigenvalue and Eigenvectors

1. $Ax = \lambda x\ for\ some\ non-zero\ x \Rightarrow \lambda\ is\ an\ eigenvalue, x\ is\ an\ eigenvector$
2. $x\ is\ an\ eigenvector\ \Leftrightarrow x \in N(A - \lambda I)\ \Leftrightarrow N(A - \lambda I)\ is\ non\ trivial \Leftrightarrow \det(A - \lambda I) = 0$
3. Eigenvalue and eigenvector for special matrices
   a. Triangle matrix: eigenvalues = diagonal
   b. Singular matrix: $\lambda = 0$ is an eigenvalue
   c. Projection matrix: for any vector in plane, $\lambda = 1$; for any vector $\perp$ plane, $\lambda = 0$ (if such vector exists)
   d. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \lambda = 0, x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow only\ 1\ engenvector$

e. Markov matrix (each column adds to 1): $\lambda = 1$ is an eigenvalue. (proof: A-I is singular)

4. $\sum \lambda = trace(A), \prod \lambda = \det(A)$
5. $A\ and\ A^T\ have\ the\ same\ \lambda s$
6. A and B share the same n independent eigenvectors if and only if AB = BA

Lecture 22 Diagonalization

1. Assume A is an $n \times n$ matrix, with k eigenvectors ($k \leq n$). Then we can write
$$AS = S\Lambda$$
where S is the $n \times k$ eigenvector matrix, Λ is the $k \times k$ eigenvalue diagonal matrix.
2. If A has n independent eigenvectors, then S is invertible, and A is diagonalizable.
$$A = S\Lambda S^{-1}, \qquad \Lambda = S^{-1}AS$$
3. A has n eigenvalues if we count the multiplicity.
4. If all λ are different, then A is sure to have n independent eigenvectors and is diagonalizable. If there are repeated λ, then we need to consider the algebraic multiplicity (AM: repetitions of λ) and geometric multiplicity (GM: number of independent eigenvectors for λ).
   - If GM=AM for all λ, then A is diagonalizable.
   - If GM<AM for some λ, then A is not diagonalizable.
5. There is no connection between invertibility and diagonalizability.
   - Invertible: $eigenvalue\ \lambda \neq 0$
   - Diagonalizable: $n\ indepedent\ eigenvectors$
6. Calculate $A^k u_0$ for some diagonalizable A
Let $u_0 = c_1x_1 + c_2x_2 + \cdots + c_nx_n = Sc$, then $A^k u_0 = A^k Sc = S\Lambda^k S^{-1}Sc = S\Lambda^k c$
7. Fibonacci: $F_{k+1} = F_k + F_{k-1}$
2nd order difference → convert to 1st order with vector
Let $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$, then $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} F_k + F_{k-1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix} = Au_{k-1}$

Lecture 23 Differential Equations

1. Solve $\frac{du}{dt} = Au$, where A is a constant matrix. For simplicity we only consider when A is diagonalizable.
$$\frac{du}{dt} = Au \Leftrightarrow \frac{du}{dt} = S\Lambda S^{-1}u \Leftrightarrow \frac{dS^{-1}u}{dt} = \Lambda S^{-1}u \Leftrightarrow S^{-1}u(t) = e^{\Lambda t}S^{-1}u(0)$$
$$\Leftrightarrow u(t) = Se^{\Lambda t}S^{-1}u(0) = e^{At}u(0)$$
2. What is e^A?
For diagonal matrix Λ,
$$e^\Lambda = I + \Lambda + \frac{\Lambda^2}{2!} + \frac{\Lambda^3}{3!} + \cdots = \begin{bmatrix} 1 + \lambda_1 + \frac{\lambda_1^2}{2!} + \cdots & & \\ & \ddots & \\ & & 1 + \lambda_n + \frac{\lambda_n^2}{2!} + \cdots \end{bmatrix} = \begin{bmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{bmatrix}$$
For $A = S\Lambda S^{-1}$,
$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots = I + S\Lambda S^{-1} + \frac{1}{2!}S\Lambda^2S^{-1} + \frac{1}{3!}S\Lambda^3S^{-1} + \cdots$$

$$= S\left(I + \Lambda + \frac{\Lambda^2}{2!} + \frac{\Lambda^3}{3!} + \cdots\right) S^{-1} = Se^{\Lambda}S^{-1}$$

3. From
$$e^{\mathrm{A}} = Se^{\Lambda}S^{-1}$$
we know:
   - e^{A} has the same eigenvector matrix as A
   - The eigenvalues of e^{A} are exponential of eigenvalue of A

Lecture 24 Markov Matrix and Fourier Series

1. Markov matrix: (1) all entries ≥ 0; (2) all columns add to 1
$$\begin{bmatrix} .1 & .01 & .3 \\ .2 & .99 & .3 \\ .7 & 0 & .4 \end{bmatrix}$$
2. Properties of a Markov matrix:
   a. $\lambda_1 = 1$ is an eigenvalue. Proof: $\det(A - I) = 0$.
   b. The eigenvector x_1 for $\lambda_1 = 1$ has all components ≥ 0
   c. all other eigenvalues $|\lambda_i| < 1$. Proof of $\lambda \leq 1$: $A^T x$ has each component as a weighted average of x, and thus cannot be larger than the largest element of x. Therefore, $\lambda \leq 1$
   Proof incomplete.
3. Orthonormal basis
$$\vec{v} = x_1\overrightarrow{q_1} + \cdots + x_n\overrightarrow{q_n} = Q\vec{x}$$
$$\overrightarrow{q_1}^T\vec{v} = x_1\overrightarrow{q_1}^T\overrightarrow{q_1} = x_1, \qquad \therefore x_1 = \overrightarrow{q_1}^T\vec{v}, \qquad \therefore \vec{x} = Q^{-1}\vec{v} = Q^T\vec{v}$$
4. Fourier series
$$f(x) = a_0 + a_1\cos(x) + b_1\sin(x) + a_2\cos(2x) + b_2\sin(2x) + \cdots$$
   a. sin and cos are orthogonal
   b. Basis: $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots$
   c. Inner product of vectors: $v^T w = v_1 w_1 + \cdots + v_n w_n$
   Inner product of functions should be like $f^T g = f(x_1)g(x_1) + \cdots + f(x_n)g(x_n)$
$$f^T g = \int_0^{2\pi} f(x)g(x)dx$$
   d. $x_1 = \overrightarrow{q_1}^T\vec{v} \rightarrow$ similarly, $a_1 = (\cos(x))^T \cdot f(x) = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(x)\,dx$
   We need to divide by π to make norm 1

Lecture 25 Review 2

1. Calculate the determinant of the matrix with $A_{ij} = \begin{cases} 1, & if\ |i-j| \leq 1 \\ 0, & otherwise \end{cases}$
   Hint: prove $D_n = D_{n-1} - D_{n-2}$
2. Calculate the determinant of matrix with $A_{ij} = \begin{cases} \min(i,j), & if\ |i-j| = 1 \\ 0, & otherwise \end{cases}$
   Hint: prove $D_{n+2} = -(n+1)^2 D_n$

Lecture 26 Symmetric Matrix, Positive Definite

1. Real symmetric matrix has following 2 properties:
    a. All eigenvalues are real
    b. The eigenvectors can be chosen orthonormal (and has n eigenvectors and thus diagonalizable)
2. Proof of properties in 1:
    a. Proof of 1a: If λ is complex, we have $Ax = \lambda x$, then $A\bar{x} = \bar{\lambda}\bar{x}$. Then $\bar{x}^T A^T = \bar{x}^T \bar{\lambda}$. As $A^T = A$, then $\bar{x}^T A = \bar{x}^T \bar{\lambda}$. Thus $\bar{x}^T A x = \bar{x}^T \bar{\lambda} x = x^T \lambda x$, so $\bar{\lambda}\bar{x}^T x = \lambda \bar{x}^T x$. As $\bar{x}^T x > 0$, therefore $\lambda = \bar{\lambda}$, and λ is real.
    b. Proof of 1b:
    When no repeated λ, $Ax_1 = \lambda_1 x_1, Ax_2 = \lambda_2 x_2$, and thus $x_1^T A x_2 = x_1^T \lambda_1 x_2 = x_1^T \lambda_2 x_2$, therefore $x_1^T x_2 = 0$
    When there are repeated λ,
3. Usual case diagonalization: $A = S\Lambda S^{-1}$
    Symmetric case: $A = Q\Lambda Q^{-1} = Q\Lambda Q^T$, spectral theorem
4. $A = Q\Lambda Q^T = \lambda_1 q_1 q_1^T + \cdots + \lambda_n q_n q_n^T$
    Every symmetric matrix is a combination of perpendicular projection matrices。
5. For complex matrix, if $A = \bar{A}^T$, then A has real eigenvalues and perpendicular eigenvectors.
6. For real symmetric matrix, the number of positive pivots = number of positive eigenvalues
7. Positive definite matric has the following equivalent properties:
    a. all $\lambda > 0$
    b. all pivots > 0
    c. all sub-determinants > 0

Lecture 27 Complex matrix, Fast Fourier Transform

1. Hermitian: $A^H = \overline{A^T}$
2. Complex vs real

Complex matrix		Real matrix	
Conjugate transpose	$A^H = \overline{A^T}$	Transpose	A^T
Hermitian matrix	$A^H = A$	Symmetric matrix	$A^T = A$
Inner product	$x^H y$	Inner product	$x^T y$
Length	$\sqrt{x^H x}$	Length	$\sqrt{x^T x}$
Perpendicular	$x^H y = 0$	Perpendicular	$x^T y = 0$
Unitary Matrix	$Q^H Q = I$	Orthogonal matrix	$Q^T Q = I$

3. Fourier matrix

$$F_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{bmatrix}, where\ w = e^{i\frac{2\pi}{n}}$$

4. $\frac{1}{\sqrt{n}} F_n$ is orthogonal matrix
5. Fast Fourier Transform

The operation $F_n x_n$ is $O(n^2)$, but fast Fourier transform is $O(n \cdot \log n)$

$$F_{64} = \begin{bmatrix} I & D \\ I & -D \end{bmatrix} \begin{bmatrix} F_{32} & 0 \\ 0 & F_{32} \end{bmatrix} \begin{bmatrix} 1 & & & & & & & \\ & & 1 & & & & & \\ & & & \cdots & & & & \\ & & & & & 1 & & \\ & 1 & & & & & & \\ & & & 1 & & & & \\ & & & & \cdots & & & \\ & & & & & & 1 & \end{bmatrix}$$

$$f(n) = 2f\left(\frac{n}{2}\right) + O(n) \to f(n) = O(n \cdot \log n)$$

Lecture 28 Positive Definite Matrix

1. For a 2×2 symmetric matrix, positive definite is equivalent to the following properties:
   a. $\lambda_1 > 0, \lambda_2 > 0$
   b. $a > 0,\ ac - b^2 > 0$
   c. $pivots\ a > 0, \frac{ac-b^2}{a} > 0$
   d. $x^T A x > 0\ for\ any\ x \neq 0$
2. In calculus, min $\sim \frac{d^2u}{dx^2} > 0$
   In linear algebra, for $f(x_1, \dots, x_n)$, min $\sim$ matrix of 2nd derivative is positive definite
3. For a positive definite matrix A, $x^T A x = 1$ is an ellips oid, with $A = Q \Lambda Q^T$:
   a. Eigenvectors are direction of principle axis
   b. Eigenvalues are $\frac{1}{half\ length}$
4. Some properties of positive definite matrix:
   a. If A is PD, then A^{-1} is also PD. Proof: $A = Q\Lambda Q^T \Rightarrow A^{-1} = Q\Lambda^{-1}Q^T$
   b. If A, B are PD, then $A + B$ is PD. Proof: $x^T A x > 0, x^T B x > 0 \Rightarrow x^T(A+B)x > 0$
   c. If A is a $m \times n$ matrix, then $A^T A$ is PD if A has independent columns.

Lecture 29 Similar Matrix, Jordan Form

1. $n \times n$ matrix A and B are similar if $B = M^{-1}AM\ \ for\ some\ matrix\ M$
   a. If A is diagonalizable, $S^{-1}AS = \Lambda$, then A and Λ are similar
   b. Similar matrices have same eigenvalues and same number of eigenvectors
      Proof: $Ax = \lambda x \Leftrightarrow AMM^{-1}x = \lambda x \Leftrightarrow M^{-1}AMM^{-1}x = \lambda M^{-1}x \Leftrightarrow BM^{-1}x = \lambda M^{-1}x$
2. If A has n distinct eigenvalues, then A is diagonalizable and similar to Λ
   If A has repeated eigenvalues, take 2×2 matrix with repeated $\lambda = 4$ as an example, there are 2 families of similar matrices:
   a. Only 1 matrix: $A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$. It has no other similar matrix, because $A = 4I$, and $M^{-1}AM = M^{-1}(4I)M = 4I = A$.
   b. Big "family" similar to $B = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$. Not diagonalizable with only 1 eigenvector.

3. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ $is\ not\ similar\ to$ $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ although they have the same eigenvalue and same number of eigenvectors, because they have different Jordan form.
4. Jordan blocks $J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$, with the same eigenvalue on diagonal, and 1 on $J_i(k, k+1)$. Each Jordan block has only 1 eigenvector.
5. Every square matrix A is similar to a Jordan matrix

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_d \end{bmatrix}, where\ \#\ of\ blocks = \#\ of\ eigenvectors$$

Symmetric matrix: eigenvectors can be made perp; like a real number

Skew-symmetric matrix: $A^T = -A$. Eigenvalues are pure imaginary

Orthogonal matrix: eigenvalues complex with $|\lambda| = 1$