Basic notation and background

- $M_n(\mathbb{F})$, $M_{m,n}(\mathbb{F})$ are the set of $n \times n$ and $m \times n$ matrices over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ (or a general field).
- Basic operations of complex numbers are assumed.
  $z = x + iy = \rho e^{i\theta}$, $\rho = |z| = |\bar{z}z|^{1/2} = \sqrt{x^2 + y^2}$, $\bar{z} = x - iy$, $z_1 + z_2, z_1 z_2, z_1/z_2$ if $z_2 \neq 0$.
- $\mathbb{F}^n$ is the set of column vectors of length n with entries in $\mathbb{F}$.
- If $\mathbb{F}$ is clear, we use the notation $M_n, M_{m,n}$.
- Let $A \in M_{m,n}$. Then $A^T \in M_{n,m}$. For complex matrix A, we have $\bar{A}$ and $A^* = (\bar{A})^T = \overline{A^T}$.
- An $m \times n$ matrix $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$.

I.1 A close look at $Ax = b$

- Linear equations, elementary row operations, solution sets.

  Recall. Let $A \in M_{m,n}(\mathbb{F}), b \in \mathbb{F}^n$, $Ax = b$.

  Find $E_r \cdots E_1[A|b] = [\tilde{A}|\tilde{b}]$ in row echelon form, $E_r \cdots E_1[A|I_n] = [I_n|A^{-1}]$.

- Matrices, column space, row space, null space, ranks.

  Recall. Let $A \in M_{m,n}(\mathbb{F})$, and $E_r \cdots E_1 A = [\tilde{A}]$ in echelon form.

  Then we can find the bases for column space, row space, and null space, and the rank of A.

Interpretation of $Ax = b$.

Example Let $A = [A_1 A_2] \in M_{3,2}, x = (x_1, x_2)^T, b \in \mathbb{F}^3$.

Then $Ax = b$ means $b = x_1 A_1 + x_2 A_2$.

All combination of A_1, A_2 form the column space.

The equation $Ax = b$ is solvable means that b is in the column space.

In general, if $A \in M_{3,n}$, the column space can be of dimensions $0, 1, 2, 3$.

All these comments hold for the general case: $Ax = b$.

For example, if $A = [A_1| \cdots |A_n] \in M_{m,n}(\mathbb{F})$ and $Ax = b \in \mathbb{F}^m$ then $b = x_1 A_1 + \cdots + x_n A_n$.

If $Ax = 0 \in \mathbb{F}^m$ has non-trivial solution then $\{A_1, \ldots, A_n\}$ is linearly independent, i.e., there is $x_1, \ldots, x_n$ not all zero such that $x_1 A_1 + \cdots + A_n = 0$.

The null space of $A \in M_{m,n}$ is the set/subspace $\{x \in \mathbb{F}^n : Ax = 0\}$.

Theorem rank + null space dimension $= n$.

Proposition If $A \in M_{m,n}$ has rank r, then $A = CR$, where $C \in M_{m,r}$ with independent columns forming a basis for the column space, and $R \in M_{r,n}$ has independent rows forming a basis for the row space. So, the row rank and column rank of A are the same.

Remark The result is useful for low rank factorization.

There will be better factorization, namely, the singular value decomposition.

I.2 Matrix Multiplication

- Note that $A \in M_{m,n}$ is rank one if and only if $A = uv^T$ for some nonzero vector $u \in \mathbb{F}^m, v \in \mathbb{F}^n$.
- In general if $A \in M_{m,n}$ has columns $u_1 \dots u_n$ and $B \in M_{n,p}$ has rows $b_1^T, \dots, b_n^T$, i.e., $A = [u_1 \cdots u_n], B^T = [b_1 \cdots b_n]$, then $AB = u_1 v_1^T + \cdots + u_n v_n^T$.

  One can check the (i, j) entry of $C = AB$ and $\sum_{j=1}^m u_i v_j^T$.
- Furthermore, if $D = \text{diag}\,(d_1, \dots, d_n)$, then $ADB = d_1 u_1 v_1^T + \cdots + d_n u_n v_n^T$.

Example

Proposition If $A \in M_{m,n}$ has rows $x_1^T, \dots, x_m^T$ and $B \in M_{n,p}$ has columns $y_1, \dots, y_n$, then

$$AB = [Ay_1 | \cdots | Ay_n] = \begin{bmatrix} x_1^T B \\ \cdots \\ x_m^T B \end{bmatrix}.$$

Consequently, if $B_1, B_2 \in M_{m,n}$ then $A(B_1 + B_2) = AB_1 + AB_2$.

Proposition. Suppose $A \in M_{m,n}, B \in M_{n,p}, C \in M_{p,q}$. Show that $(AB)C = A(BC)$.

Case 1. Suppose $B = xy^T$ with $\mathbf{x} = (x_1, \dots, x_n)^t$, $\mathbf{y}^T = (y_1, \dots, y_p)$.

Let $A\mathbf{x} = (\hat{x}_1, \dots, \hat{x}_m)^T$ and $\hat{\mathbf{y}}^T = \mathbf{y}^T C = (\hat{y}_1, \dots, \hat{y}_q)$. Then

$$\begin{aligned}(AB)C &= (Axy^T)C = (A[\mathbf{x}y_1 | \cdots | \mathbf{x}y_p])C = [(A\mathbf{x})y_1 | \cdots | (A\mathbf{x})y_p]C = ((A\mathbf{x})\mathbf{y}^T)C \\ &= (\hat{\mathbf{x}}\mathbf{y}^T)C = \begin{bmatrix} \hat{x}_1 \mathbf{y}^T \\ \vdots \\ \hat{x}_m \mathbf{y}^T \end{bmatrix} C = \begin{bmatrix} \hat{x}_1 \mathbf{y}^T C \\ \vdots \\ \hat{x}_m \mathbf{y}^T C \end{bmatrix} = \begin{bmatrix} \hat{x}_1 \hat{\mathbf{y}}^T \\ \vdots \\ \hat{x}_m \hat{\mathbf{y}}^T \end{bmatrix} = \hat{x}\hat{y}^T.\end{aligned}$$

Similarly,

$$A(BC) = A(\mathbf{x}\mathbf{y}^T C) = A \begin{bmatrix} x_1 \hat{\mathbf{y}}^T \\ \cdots \\ x_n \mathbf{y}^T \end{bmatrix} = [(A\mathbf{x})\hat{y}_1 | \dots | (A\mathbf{x})\hat{y}_q] = (\hat{\mathbf{x}})\hat{y}^T.$$

Case 2. If $B = \sum_{j=1}^r \mathbf{x}_j \mathbf{y}_j^T$, then

$$(A(\sum_{j=1}^r \mathbf{x}_j \mathbf{y}_j^T))C = (\sum_{j=1}^r A(\mathbf{x}_j \mathbf{y}_j^T))C = \sum_{j=1}^r A((\mathbf{x}_j \mathbf{y}_j^T)C)) = A(\sum_{j=1}^r \mathbf{x}_j \mathbf{y}_j^T C) = A(BC).$$

Matrix factorization.

$$A = LU, \quad A = QR, \quad A = RDR^{-1}S \quad A = Q\Lambda Q^T \text{ if } A \text{ is symmetric}, \quad A = U\Sigma V^T.$$

1. LU factorization for invertible $A \in M_n$.

   To solve $Ax = b$ for many different vectors $b \in \mathbb{F}^n$.

   Write $A = LU$ where L is lower triangular, and U as upper triangular.

   Then solve $Ax = LUx = b$ by solving $Ly = b$ and $Ux = y$.

   How to write $LU = A$? We will see this in Section 1.4

   One can do an exchange of rows by a permutation matrix P to get $PA = LU$ if A is invertible.

   Question What if A is not invertible or A is not a square matrix?

2. QR Decomposition for $A \in M_{m,n}$ with $m \geq n$.

   Recall the inner product on $\mathbb{F}^n$ defined by $\langle x, y\rangle = y^*x = \sum_{j=1}^n x_j\bar{y}_j$, and the inner product norm $\|x\| = \langle x, x\rangle^{1/2}$.

   We are interested in orthonormal set/basis $\{v_1, \ldots, v_n\}$ such that $V^*V = I_n$ if $V = [v_1|\cdots|v_n]$, i.e., $\langle v_i, v_j\rangle = \delta_{ij}$.

   To solve $Ax = b$ with b in the column space of b, we can solve $Rx = Q^*Ax = Q^*b$.

3. $A = RDR^{-1} = \sum_{j=1}^n \lambda_j x_j y_j^T$ where R has columns $x_1, \ldots, x_n$ and R^{-1} has rows $y_1^T, \ldots, y_n^T$. Then $A^k = \sum_{j=1}^n \lambda_j^k x_j y_j^T$.

   If $H = H^*$ is Hermitian, then $H = Q\Lambda Q^* = \sum_{j=1}^n \lambda_j v_j v_j^*$ with $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ so that $H^k = \sum_{j=1}^n \lambda_j^k v_j v_j^*$.

4. $A = U\Sigma V^*$. Then $A = \sum_{j=1}^r s_j u_j v_j^*$.

I.3 The Four Fundamental Subspaces

Let $A \in M_{m,n}$. We have
the column space $C(A)$ in $\mathbb{F}^m$ contains all combination of columns of A,
the null space $N(A)$ contains x in $\mathbb{F}^n$ such that $Ax = 0$.
the row space $C(A^T)$ contains all combination of columns of A^T,
the left null space $N(A^T)$ contains y in $\mathbb{F}^m$ such that $A^T y = 0$.

Proposition Let $A \in M_{m,n}$. Then $C(A)$ and $C(A^T)$ have the same dimension r; $N(A)$ has dimension $n - r$, $N(A^T)$ has dimension $m - r$.

Theorem Let $A \in M_{m,r}, B \in M_{r,n}$.

1. $\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$.
2. $\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$.
3. $\operatorname{rank}(A^*A) = \operatorname{rank}(AA^*) = \operatorname{rank}(A) = \operatorname{rank}(A^*)$.
4. $\operatorname{rank}(AB) = r$ if $\operatorname{rank}(A) = \operatorname{rank}(B) = r$.