$2x_1 + 3x_2 = b_1$
$2x_1 + 4x_2 = b_2$
$3x_1 + 7x_2 = b_3$

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

$$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

- 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, $E_r \cdots E_1[A|I_n] = [I_n|A^{-1}]$.

$$x_1\begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

- Matrices, column space, row space, null space, ranks. $\rightarrow$ rank(A) = # non-zero rows of the row echelon form of A $= \dim C(A^T) = \dim C(A)$

  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.

$A \in M_{m,n}$, null space $N(A) = \{x \in \mathbb{F}^n : Ax = 0\}$

Interpretation of $Ax = b$.

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

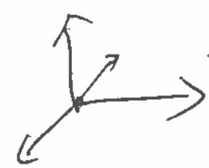
Then $Ax = b$ means $b = x_1 A_1 + x_2 A_2$.

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 $0, 1, 2, 3$.

$\mathbb{R}^3$ $\quad A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ $C(A) = \left\{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\right\}$



$A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix}$ $C(A) = 1$ is a line $= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 5 \end{bmatrix}$

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

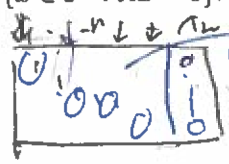
For example, if $A = [A_1|\cdots|A_n] \in M_{m,n}(\mathbb{F})$ and $Ax = b \in \mathbb{F}^m$ then $b = x_1 A_1 + \cdots + x_n A_n$.

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 $= n$.

$m \times n$ $\quad Ax = 0$

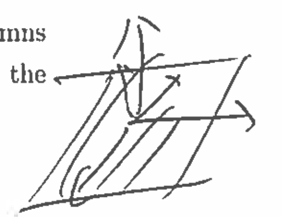


rank r. dim null(A) $= n - r$

$A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 7 \end{bmatrix}$ $C(A)$

$A \in M_{m,n}$ pivoting one + rest $= n$

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.



$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{m1} \\ a_{12} & \vdots & \vdots \\ a_{1n} & a_{2n} & a_{mn} \end{bmatrix}$$

Remark The result is useful for low rank factorization.

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

rank(A) $= r$

rank(A^T) $= r$

$(XY)^T = Y^T X^T$ $\quad A^T y = 0, \; y^T A = 0$

dim null(A^T) $= m - r$

Example

$$\begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \end{bmatrix}$$

Example:

$$\begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}$$

Example

$$\begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \underset{m \times r}{\begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}} \underset{r \times n}{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}}$$

$$\begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} + 3 \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \\ 2 \end{bmatrix}$$

Remark: ① $\underset{m \times n}{A} \underset{n \times p}{[B_1 | \cdots | B_p]} = [AB_1 | \cdots | AB_p]$

② $\begin{bmatrix} A_1^T \\ \vdots \\ A_m^T \end{bmatrix} B = \begin{bmatrix} A_1^T B \\ \vdots \\ A_m^T B \end{bmatrix}$

---

③ $A = [x_1 | \cdots | x_n] \quad B = \begin{bmatrix} y_1^T \\ \vdots \\ y_n^T \end{bmatrix}$

$$AB = [x_1 | \cdots | x_n] \begin{bmatrix} y_1^T \\ \vdots \\ y_n^T \end{bmatrix}$$

$$= \underbrace{\underset{m \times 1}{x_1} \underset{1 \times p}{y_1^T}}_{m \times p} + \cdots + \underbrace{x_n y_n^T}_{m \times p}$$

Proof: Check the (i,j) entry on both side. □

Example:

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

$$= \begin{bmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \\ \cdots \end{bmatrix} + \begin{bmatrix} 0 & 3 & 6 \\ 0 & 2 & 4 \\ \cdots \end{bmatrix} = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ \cdots \end{bmatrix}$$

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 $D = \operatorname{diag}(d_1, \dots, d_n)$, then $ADB = d_1 u_1 v_1^T + \cdots + d_n u_n v_n^T$.

$m \times n \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} n \times p$

$$[u_1 \cdots u_n]\begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}\begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} = [\quad]\begin{bmatrix} d_1 v_1^T \\ d_2 v_2^T \\ \vdots \\ d_n v_n^T \end{bmatrix}$$

$$= u_1 d_1 v_1^T + u_2 d_2 v_2^T + \cdots + u_n d_n v_n^T = d_1 u_1 v_1^T + \cdots + d_n u_n v_n^T.$$

Example

$$A = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

$= $ [illegible] $= \hat{A}\begin{bmatrix} \sqrt{10} & 0 \\ 0 & \sqrt{70} \end{bmatrix}\hat{B}$

$= \sqrt{10}\, \hat{u}_1 \hat{v}_1^T + \sqrt{70}\, \hat{u}_2 \hat{v}_2^T$

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

$B_1 = [v_1 | \cdots | v_p] \qquad B_2 = [w_1 | \cdots | w_p]$

$A(B_1 + B_2) = A[v_1 + w_1 | \cdots | v_p + w_p] = [Av_1 + Aw_1 | \cdots | Av_n + Aw_p]$

$= [Av_1 | \cdots | Av_n] + [Aw_1 | \cdots | Aw_p]$

$= AB_1 + AB_2.$

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

$m \times p \; p \times q \qquad m \times n \; n \times q$

Case 1 $\quad B = x y^T \qquad x \in \mathbb{F}^{n \times 1} = (x_1 \cdots x_n)^T \qquad y^T \in \mathbb{F}^{1 \times p} = (y_1 \cdots y_p)$

$(AB)C = (A x y^T) C = (A [x y_1 | x y_2 | \cdots | x y_p]) C$

$= [A x y_1 | \cdots | A x y_p] C = [\hat{x} y_1 | \cdots | \hat{x} y_p] C \qquad \hat{x} = Ax = \begin{bmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_m \end{bmatrix}$ $m \times 1$

$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_n \end{bmatrix} [y_1 \cdots y_p]$

$= \begin{bmatrix} \hat{x}_1 y^T \\ \hat{x}_2 y^T \\ \vdots \\ \hat{x}_m y^T \end{bmatrix} C = \begin{bmatrix} \hat{x}_1 (y^T C) \\ \vdots \\ \hat{x}_m (y^T C) \end{bmatrix} = \hat{x}\, \hat{y}^T$

$\hat{y}^T = yC = [\hat{y}_1, \cdots, \hat{y}_q]$

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

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

$(Ax)(y^TC)$

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

$[A(B_1 + \cdots + B_r)]C = [(AB_1 + \cdots + AB_r)]C$
$= (AB)C + \cdots + (AB)C$

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

$= A(B_1C) + \cdots A(B_rC$
$= A(BC)$

$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ x & x \end{bmatrix} \begin{bmatrix} 0 & x \\ 0 & x \end{bmatrix}$

Matrix factorization.

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

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

   To solve $Ax = b$ for many different vectors $b \in \mathbb{F}^n$. $Ax = b_1, \; Ax = b_2, \; \ldots \; Ax = b_\ell \; \ldots$

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

   Then solve $\boxed{Ax = LUx = b \text{ by solving } Ly = b \text{ and } Ux = y.}$

   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 $\boxed{PA = LU}$ 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., $\langle v_i, v_j\rangle = \delta_{ij}$. V^T real; Complex

   $Q^TQ = I_n$

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

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 $A^k = \sum_{j=1}^n \lambda_j^k x_j y_j^T$. $A = [R] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} R^{-1}$

   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 $A = \sum_{j=1}^r s_j u_j v_j^*$.

Remark: Solving $Ax = b$ by $x = A^{-1}b$ is not a good way

because A^{-1} may be bad, A is ill-conditioned.

Example $\begin{bmatrix} 1 & 1 \\ 1 & 1.0000001 \end{bmatrix}$ $Dx = b$

$A = \underbrace{\begin{bmatrix} 1 & 0 \\ * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} * & * \\ 0 & 1 \end{bmatrix}}_{U}$ $\begin{bmatrix} d_1 & \\ 0 & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$