Note Title 2025-02-27

To be confirmed: proposal: move lecture
Fri 16-18 $\to$ Wed 16-18

Vectorization: basis for the vector space $\mathbb{C}^{m\times n}$

$$\mathrm{vec}: \mathbb{C}^{m\times n} \to \mathbb{C}^{mn}$$

$$X = \left[\begin{array}{c|c|c|c} x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & & x_{mn} \end{array}\right] = \begin{bmatrix} x_{11} \\ \vdots \\ x_{m1} \\ \hline x_{12} \\ \vdots \\ x_{m2} \\ \vdots \\ \hline x_{1n} \\ \vdots \\ x_{mn} \end{bmatrix}$$

"column-major order"

```
c = zeros
for i = 1:n
    for j = 1:n
        c[i] = c[i] + A[i,j] b[j]
end
end
```

$c = Ab$

swap to access matrix in column-major

$$(\mathrm{vec}\, X)_{i+m(j-1)} = X_{ij}$$ if columns are indexed from 1

$i + mj$ if indexing from 0

$X \mapsto AXB$ is a linear map for X (given fixed A, B of suitable size)

$m\times n$ $\quad$ $p\times q$

There is a matrix $K \in \mathbb{C}^{pq \times mn}$ such that

$$\operatorname{vec}(AXB) = K \cdot \operatorname{vec}(X) \quad \text{for all } X \in \mathbb{C}^{m \times n}$$

$$pq \qquad pq\,\underset{mn}{\square} \qquad mn$$

$$K = \begin{bmatrix} b_{11}A & b_{21}A & b_{31}A & \cdots b_{n1}A \\ b_{12}A & A & A & A \\ \vdots & & & \\ b_{1q}A & A & A & \cdots b_{nq}A \end{bmatrix} = B^T \otimes A$$

<u>Kronecker product</u> of B^T and A.

<u>Def</u>: Given $F \in \mathbb{C}^{mn}$, $G \in \mathbb{C}^{pq}$,

$F \otimes G$ is a $mp \times nq$ matrix with blocks

$$F \otimes G \quad \begin{bmatrix} f_{11}G & f_{12}G & f_{13}G & \cdots & f_{1n}G \\ \vdots & & & & \\ f_{m1}G & f_{m2}G & & \cdots & f_{mn}G \end{bmatrix}.$$

<u>Properties</u>:

- $\operatorname{vec}(AXB) = (B^T \otimes A) \cdot \operatorname{vec} X$ ⚠ <u>not</u> transpose-conjugate!
- $(\alpha A + \beta B) \otimes C = \alpha (A \otimes C) + \beta (B \otimes C)$
  $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
- $(A \otimes B)^T = A^T \otimes B^T$
- $(A \otimes B) \cdot (C \otimes D) = (AC \otimes BD)$ when the dimensions allow it

Proof: for any $x = \mathrm{vec}(X)$, we have

$$(A\otimes B)(C\otimes D)\,\mathrm{vec}\,X = (A\otimes B)\,\mathrm{vec}(DXC^T) = \mathrm{vec}\,B(DXC^T)A^T$$
$$= \mathrm{vec}\,(BD)X(AC)^T = (AC\otimes BD)\,\mathrm{vec}\,X$$

Sylvester equation: given $A\in\mathbb{C}^{m\times m}$, $B\in\mathbb{C}^{n\times n}$ and $C\in\mathbb{C}^{m\times n}$, find $X\in\mathbb{C}^{m\times n}$, s.t.

$$AX - XB = C$$

This is a $mn\times mn$ linear system in the entries of X.

In matrix form:

$$\mathrm{vec}(C) = \mathrm{vec}(AX - XB) = \mathrm{vec}(AXI) - \mathrm{vec}(XB)$$
$$= (I\otimes A)\,\mathrm{vec}\,X - (B^T\otimes I)\,\mathrm{vec}\,X$$

$$\overbrace{(I\otimes A - B^T\otimes I)}^{K}(\mathrm{vec}\,X) = \mathrm{vec}\,C$$

Theorem: The Sylvester equation $AX-XB=C$ has a unique solution if and only if the spectra of A and B are disjoint, i.e.

$$\Lambda(A)\cap\Lambda(B) = \phi$$

Proof: the eq. has a unique solution iff K is nonsingular.

Let us take Schur decompositions

$A = Q_A U_A Q_A^*$, $\qquad B^T = Q_B U_B Q_B^*$

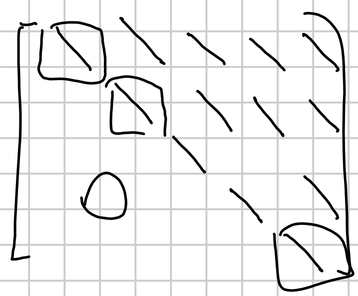
$\mathrm{diag}(U_A) = \Lambda(A)$ with alg. multiplicities $= (\lambda_1, \lambda_2 \dots \lambda_m)$

$\mathrm{diag}(U_B) = \Lambda(B) = (\mu_1, \mu_2, \dots \mu_n)$

$$K = I \otimes A - B^T \otimes I = \underbrace{(Q_B \otimes Q_A)}_{\text{orthogonal}} \underbrace{(I \otimes U_A - U_B \otimes I)}_{\text{upper triangular}} \underbrace{(Q_B^* \otimes Q_A^*)}_{\text{orth}^*}$$

This is a Schur decomposition!

$$(Q_B \otimes Q_A)(Q_B \otimes Q_A)^* = (Q_B \otimes Q_A)(Q_B^* \otimes Q_A^*) = I_n \otimes I_m = I_{mn}$$

$I \otimes U_A =$  $\qquad U_B \otimes I =$ 

$\mathrm{diag}(I \otimes U_A) = (\lambda_1, \lambda_2, \dots \lambda_m, \lambda_1, \lambda_2, \dots \lambda_m, \dots \lambda_1, \lambda_2, \dots \lambda_m)$

$\mathrm{diag}(U_B \otimes I) = (\mu_1, \mu_1, \dots \mu_1, \mu_2 \mu_2 \dots \mu_2, \dots, \mu_n, \mu_n, \dots \mu_n)$

$\mathrm{diag}(K) = (\lambda_i - \mu_j, \; i = 1, \dots m, \; j = 1, \dots n)$ in a suitable order

K invertible $\Leftrightarrow$ no zero eigenvalues $\Leftrightarrow$ $\lambda_i \neq \mu_j$ for all i, j. $\square$

Another instance of the same trick:

$$A \otimes B = (U_A S_A V_A^*) \otimes (U_B S_B V_B^*) = \underbrace{(U_A \otimes U_B)}_{\text{orthogonal}} \underbrace{(S_A \otimes S_B)}_{\text{diagonal}} \underbrace{(V_A \otimes V_B)^*}_{\text{orth.}}$$

is (up to reordering) a singular value decomposition of $A \otimes B$ with diagonal

$$\left(\sigma_i^A \cdot \sigma_j^B \ , \ i=1,\dots m, \ j=1,\dots n \right)$$

Lemma: the sing. values of $A \otimes B$ are the pairwise products of the sing. values of A and of B.

Corollary: $\|A \otimes B\| = \|A\| \otimes \|B\|$ for the Euclidean norm.

How to solve Sylvester equations numerically?

$$(I \otimes A - B^T \otimes I) \operatorname{vec}(X) = \operatorname{vec} C$$

$mn \times mn$ linear system $\to$ Gaussian elimination very slow $O(m^3 n^3)$

Bartels-Stewart algorithm (1972)

Idea:

$$K = (Q_B \otimes Q_A)(I \otimes U_A - U_B \otimes I)(Q_B \otimes Q_A)^*$$

Practice:

$$AX - XB = C$$

$$Q_A U_A Q_A^* X - X \overline{Q_B} U_B^T Q_B^T = C$$

Multiply by Q_A^* and $\overline{Q_B}$:

$$\cancel{Q_A^* Q_A} U_A \underbrace{Q_A^* X \overline{Q_B}}_{Y} - \underbrace{Q_A^* X \overline{Q_B}}_{Y} U_B^T \cancel{Q_B^T \overline{Q_B}} = \underbrace{Q_A^* C \overline{Q_B}}_{E}$$

$$U_A Y - Y U_B^T = E \quad \text{Sylv. equation with}$$

triangular coefficients!

We can solve this new Sylvester equation via a backsubstitution:

$$\underbrace{\begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}}_{U_A} \underbrace{\begin{bmatrix} \times & \times & \times \\ \times & \boxed{\times} & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}}_{Y} - \underbrace{\begin{bmatrix} \times & \times & \times \\ \times & \boxed{\times} & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}}_{Y} \underbrace{\begin{bmatrix} \times & 0 & 0 \\ \times & \times & 0 \\ \times & \times & \times \end{bmatrix}}_{U_B^T} = \underbrace{\begin{bmatrix} \times & \times & \times \\ \times & \boxed{\times} & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}}_{E}$$

The equation corresponding to position (i,j) involves only Y_{ij} and entries <u>below</u> and to the <u>right</u> of it

$$E_{ij} = (U_A Y - Y U_B^T)_{ij} = \sum_{k \geq i} (U_A)_{ik} Y_{kj} - \sum_{k \geq j} Y_{ik} (U_B^T)_{kj}$$

$$= \boxed{(U_A)_{ii} Y_{ij}} + (U_A)_{i,i+1} Y_{i+1,j} + \dots + (U_A)_{im} Y_{mj} - \left(\boxed{Y_{ij} {U_B^T}_{jj}} + Y_{i,j+1} (U_B^T)_{j+1,j} + \dots + Y_{in} (U_B^T)_{nj} \right)$$

We can start from the bottom-right entry equation

$$(U_A)_{mm} Y_{mn} - Y_{mn} (U_B^T)_{nn} = E_{mn}$$

and solve it for $Y_{mn} = \dfrac{E_{mn}}{(U_A)_{mm} - (U_B^T)_{nn}}$

Then, each equation can be solved for Y_{ij} as long as we have already computed all entries of Y to the <u>right</u> and <u>below</u> (i,j)

We can solve for the entries of Y in this order

$$Y = \begin{bmatrix} 12 & 8 & 4 \\ 11 & 7 & 3 \\ 10 & 6 & 2 \\ 9 & 5 & 1 \end{bmatrix}$$

Bartels-Stewart Algorithm

1. Compute Schur decompositions $A = Q_A U_A Q_A^*$ $\quad O(m^3)$

   $B^T = Q_B U_B Q_B^*$ $\quad O(n^3)$

2. Compute $E = Q_A^* C \overline{Q_B}$ $\quad O(m^2 n)$

3. Solve by "back-substitution"

   $$U_A Y - Y U_B^T = E \qquad O(mn(m+n))$$

   to compute Y

4. Compute $X = Q_A Y Q_B^T$ $\quad O(m^2 n)$

steps 2-3-4 correspond to inverting factor-by-factor the Schur decomposition of K:

$$\operatorname{vec} X = K^{-1} \operatorname{vec} C = \underbrace{(Q_B \otimes Q_A)\underbrace{(I \otimes U_A - U_B^T \otimes I)^{-1}\underbrace{(Q_B \otimes Q_A)^* \operatorname{vec} C}_{2}}_{3}}_{4}$$

$$\left((U_A)_{ii} - (L_B)_{jj}\right) \boxed{Y_{ij}} = E_{ij} - \sum_{k>i} (U_A)_{ik} Y_{kj} + \sum_{k>j} Y_{ik} (U_B^T)_{kj}$$