Sylvester equations

Goal represent linear functions $\mathbb{R}^{m \times n} \to \mathbb{R}^{p \times q}$.

For instance, to deal with problems like the following one.

Sylvester equation

$$AX - XB = C$$

$A \in \mathbb{C}^{m \times m}$, $C, X \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times n}$.

This must be a $mn \times mn$ linear system, right?

Vectorization gives an explicit way to map it to a vector.

Vectorization: definition

$$\operatorname{vec} X = \operatorname{vec}\left[\begin{array}{c|c|c|c} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{array}\right] := \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \\ \hline x_{12} \\ x_{22} \\ \vdots \\ x_{m2} \\ \hline \vdots \\ \hline x_{1n} \\ x_{2n} \\ \vdots \\ x_{mn} \end{bmatrix}.$$

$$\mathbb{R}^{m\times n} \to \mathbb{R}^{mn}$$

Vectorization: comments

Column-major order: leftmost index ‘changes more often’. Matches Fortran, Matlab standard (C/C++ prefer row-major instead).

Converting indices in the matrix into indices in the vector:

$$(X)_{ij} = (\operatorname{vec} X)_{i+mj} \qquad \text{0-based,}$$
$$(X)_{ij} = (\operatorname{vec} X)_{i+m(j-1)} \qquad \text{1-based.}$$

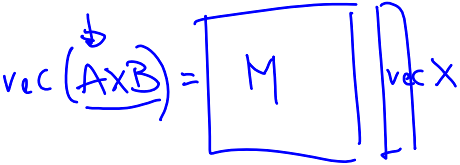
vec(AXB)

First, we will work out the representation of a simple linear map, $X \mapsto AXB$ (for fixed matrices A, B of compatible dimensions).

If $X \in \mathbb{R}^{m \times n}$, $AXB \in \mathbb{R}^{p \times q}$, we need the $pq \times mn$ matrix that maps $\operatorname{vec} X$ to $\operatorname{vec}(AXB)$.

$$(AXB)_{hl} = \sum_j (AX)_{hj}(B)_{jl} = \sum_j \sum_i A_{hi} X_{ij} B_{jl}$$

$$= \left[\begin{array}{cccc|cccc|c} A_{h1}B_{1l} & A_{h2}B_{1l} & \dots & A_{hm}B_{1l} & A_{h1}B_{2l} & A_{h2}B_{2l} & \dots & A_{hm}B_{2l} & \dots \end{array}\right.$$
$$\left.\begin{array}{|ccc} A_{h1}B_{nl} & A_{h2}B_{nl} & A_{hm}B_{nl} \end{array}\right] \operatorname{vec} X$$



$$\sum_{h,k} A_{1h} X_{hk} B_{k1}$$

Kronecker product: definition

$$\operatorname{vec}(AXB) = \begin{bmatrix} b_{11}A & b_{21}A & \dots & b_{n1}A \\ b_{12}A & b_{22}A & \dots & b_{n2}A \\ \vdots & \vdots & \ddots & \vdots \\ b_{1q}A & b_{2q}A & \dots & b_{nq}A \end{bmatrix} \operatorname{vec} X$$

Each block is a multiple of A, with coefficient given by the corresponding entry of $B^\top$.

Definition

$$X \otimes Y := \begin{bmatrix} x_{11}Y & x_{12}Y & \dots & x_{1n}Y \\ x_{21}Y & x_{22}Y & \dots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ x_{m1}Y & x_{m2}Y & \dots & x_{mn}Y \end{bmatrix}.$$

so the matrix above is $B^\top \otimes A$.

Properties of Kronecker products

$$X \otimes Y = \begin{bmatrix} x_{11}Y & x_{12}Y & \dots & x_{1n}Y \\ x_{21}Y & x_{22}Y & \dots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ x_{m1}Y & x_{m2}Y & \dots & x_{mn}Y \end{bmatrix}.$$

- $\operatorname{vec} AXB = (B^\top \otimes A) \operatorname{vec} X$. (Warning: not B^*, if complex).
- $(A \otimes B)(C \otimes D) = (AC \otimes BD)$, when dimensions are compatible. Proof: $B(DXC^\top)A^\top = (BD)X(AC)^\top$.
- $(A \otimes B)^\top = A^\top \otimes B^\top$.
- orthogonal $\otimes$ orthogonal $=$ orthogonal.
- upper triangular $\otimes$ upper triangular $=$ upper triangular.
- One can "factor out" several decompositions, e.g.,

orth. diag. orth.

$$A \otimes B = (U_1 S_1 V_1^*) \otimes (U_2 S_2 V_2^*) = (U_1 \otimes U_2)(S_1 \otimes S_2)(V_1 \otimes V_2)^*.$$

Solvability criterion

Theorem

The Sylvester equation is solvable for all C iff $\Lambda(A) \cap \Lambda(B) = \emptyset$.

$AX - XB = C \iff$

$$(I_n \otimes A - B^\top \otimes I_m)\,\mathrm{vec}(X) = \mathrm{vec}(C).$$

Schur decompositions of $A, B^\top$: $A = Q_A T_A Q_A^*$, $B^\top = Q_B T_B Q_B^*$.
Then,

$$I_n \otimes A - B^\top \otimes I_m = (Q_B \otimes Q_A)(I_n \otimes T_A + T_B \otimes I_m)(Q_B \otimes Q_A)^*.$$

is a Schur decomposition.
What is on the diagonal of $I_n \otimes T_A + T_B \otimes I_m$?
If $\Lambda(A) = \{\lambda_1, \ldots, \lambda_m\}$, $\Lambda(B) = \{\mu_1, \ldots, \mu_n\}$, then it's
$\Lambda(I_n \otimes A - B^\top \otimes I_m) = \{\lambda_i - \mu_j : i, j\}$.

Solution algorithms

The naive algorithm costs $O((mn)^3)$. One can get down to $O(m^3 n^2)$ (full steps of GMRES, for instance.)

Bartels–Stewart algorithm (1972): $O(m^3 + n^3)$.

Idea: invert factor by factor the decomposition

$$(Q_B \otimes Q_A)(I_n \otimes T_A + T_B \otimes I_m)(Q_B \otimes Q_A)^*.$$

- Solving orthogonal systems $\iff$ multiplying by their transpose, $O(m^3 + n^3)$ using the $\otimes$ structure.
- Solving upper triangular system $\iff$ back-substitution; costs $O(\text{nnz}) = O(m^3 + n^3)$.

Bartels–Stewart algorithm

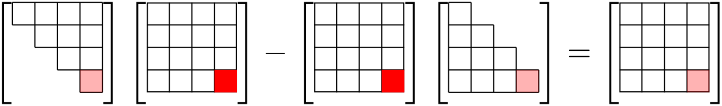
A more operational description...

Step 1: reduce to a triangular equation.

$$Q_A T_A Q_A^* X - X Q_B T_B^* Q_B^* = C$$

$$T_A \widehat{X} - \widehat{X} T_B^* = \widehat{C}, \quad \widehat{X} = Q_A^* X Q_B, \widehat{C} = \widehat{Q_A^* C Q_B}.$$

Step 2: We can compute each entry X_{ij}, by using the (i, j)th equation, as long as we have computed all the entries below and to the right of X_{ij}.



Comments

- Works also with the real Schur form: back-sub yields block equations which are tiny 2×2 or 4×4 Sylvesters.
- Backward stable (as a system of mn linear equations): it's orthogonal transformations + back-sub.
- Not backward stable in the sense of $\widetilde{A}\widetilde{X} - \widetilde{X}\widetilde{B} = \widetilde{C}$ [Higham '93].
  (Sketch of proof: backward error given by a linear least squares system with matrix $\begin{bmatrix} \widetilde{X}^\top \otimes I & I \otimes \widetilde{X} & I \end{bmatrix}$). Its singular values depend on those of $\widetilde{X}$.)

Comments

Condition number: depends on

$$\mathrm{sep}(A,B) := \sigma_{\min}(I \otimes A - B^\top \otimes I) = \min_Z \frac{\|AZ - ZB\|_F}{\|Z\|_F}.$$

If A, B normal, this is simply the minimum difference of their eigenvalues. Otherwise, it might be larger; no simple expression.

$$\sigma_{\min}(M) = \min_{v \neq 0} \frac{\|Mv\|_2}{\|v\|_2}$$

$\mathrm{sep}(A,B) = 0$ se A,B hanno autoval. comuni, >0 altrimenti

se A,B normali: T_A, T_B diagonali

$$\sigma_{\min}(I \otimes A - B^T \otimes I) = \sigma_{\min}(I \otimes T_A - T_B \otimes I) = \min_{i,j} |(T_A)_{ii} - (T_B)_{jj}|$$

Decoupling eigenvalues

Solving a Sylvester equation means finding

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

Idea Indicates how ‘difficult’ (ill-conditioned) it is to go from block-triangular to block-diagonal. (Compare also with the scalar case / Jordan form.)

Similar problem: reordering Schur forms (swapping blocks). One uses the Q factor from the QR of $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]\ldots$

Invariant subspaces

Invariant subspace (for a matrix M): any subspace $\mathcal{U}$ such that $M\mathcal{U} \subseteq \mathcal{U}$. Completing a basis U_1 to one $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ of $\mathbb{C}^m$, we get

$$U^{-1}MU = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}.$$

$MU_1 = U_1A$. $\Lambda(A) \subseteq \Lambda(M)$.
Idea: invariant subspaces are 'the span of some eigenvectors' (usually).
Example: stable invariant subspace: x s.t. $\lim_{k\to\infty} A^k x = 0$

Examples (stable invariant subspaces)

Example 1 $\operatorname{span}(v_1, v_2, \ldots, v_k)$ (eigenvectors).

Example 2 Invariant subspaces of $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$.

Example 3 Invariant subspaces of a larger Jordan matrix.

(That's the general case — idea: find Jordan form of the linear map $\mathcal{U} \mapsto \mathcal{U}$, $x \to Mx$.)

Sensitivity of invariant subspaces

If I perturb M to $M + \delta_M$, how much does U_1 change?

Proof (sketch:)

- Suppose $U = I$ for simplicity (just a change of basis).
- $M + \delta M = \begin{bmatrix} A+\delta_A & C+\delta_C \\ \delta_D & B+\delta_B \end{bmatrix}$
- Look for a transformation $V^{-1}(M + \delta M)V$ of the form $V = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$ that zeroes out the $(2,1)$ block.
- Formulate a Riccati equation $XA - BX = \delta_D - X(C + \delta C)X - X\delta_A + \delta_B X$.
- See it as a fixed-point problem

$$X_{k+1} = T^{-1}(\delta_D - X(C + \delta C)X - X\delta_A + \delta_B X)$$

- Pass to norms, show that the map sends a ball $B(0, \rho)$ to itself: $\|X\|_F \leq \|T^{-1}\|(\dots)$. For a sufficiently small perturbation, it does.

Theorem [Stewart Sun book V.2.2]

Let $M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, $\delta_M = \begin{bmatrix} \delta_A & \delta_B \\ \delta_D & \delta_C, \end{bmatrix}$, $a = \|\delta_A\|$ and so on.
If $4(\mathrm{sep}(A, B) - a - b)^2 - d(\|C\| + c) \geq 0$, then there is a (unique) X with $\|X\| \leq 2\frac{d}{\mathrm{sep}(A,B)-a-b}$ such that $\begin{bmatrix} I \\ X \end{bmatrix}$ is an invariant subspace of $M + \delta_M$.

(Not exactly what we obtain directly from the above argument — handles a and b in a slightly different way.)

Applications of Sylvester equations

Apart from the ones we have already seen:

- As a step to compute matrix functions.
- Stability of linear dynamical systems.
  Lyapunov equations $AX + XA^\top = B$, B symmetric.
- As a step to solve more complicated matrix equations (Newton's method $\to$ linearization).

We will re-encounter them later in the course (time permitting).