EE/ACM 150 - Applications of Convex Optimization in Signal Processing and Communications
Lecture 4

Andre Tkacenko

Signal Processing Research Group
Jet Propulsion Laboratory

April 12, 2012

Caltech

Outline

1. Special Vector/Matrix Operators
   - Hadamard Product
   - Kronecker Product
   - Vectorization
2. The Schur Complement
   - Introduction
   - Role in Linear Matrix Inequalities
   - Applications in Control Theory

The Hadamard or Entrywise Product

If $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m \times n}$, then the *Hadamard product* or *entrywise product* is the matrix $\mathbf{A} \circ \mathbf{B} \in \mathbb{F}^{m \times n}$ whose elements are given as follows:

$$[\mathbf{A} \circ \mathbf{B}]_{k,\ell} \triangleq [\mathbf{A}]_{k,\ell} [\mathbf{B}]_{k,\ell} \ , \ 1 \leq k \leq m, 1 \leq \ell \leq n \, .$$

Properties:

- $\mathbf{A} \circ \mathbf{B} = \mathbf{B} \circ \mathbf{A}$.
- $\mathbf{A} \circ (\mathbf{B} \circ \mathbf{C}) = (\mathbf{A} \circ \mathbf{B}) \circ \mathbf{C}$.
- $\mathbf{A} \circ (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \circ \mathbf{B}) + (\mathbf{A} \circ \mathbf{C})$.
- If $\mathbf{C} \in \mathbb{F}^{m \times m}$, then $\mathbf{C} \circ \mathbf{I}_m = \mathrm{diag}(\mathbf{C})$.
- If $\mathbf{x} \in \mathbb{F}^m$ and $\mathbf{y} \in \mathbb{F}^n$, then $\mathbf{x}^\dagger (\mathbf{A} \circ \mathbf{B}) \mathbf{y} = \mathrm{tr}\big(\mathrm{diag}(\mathbf{x}^*) \mathbf{A} \mathrm{diag}(\mathbf{y}) \mathbf{B}^T\big)$.
- (*Schur product theorem:*) If $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m \times m}$ and $\mathbf{A}, \mathbf{B} \succeq \mathbf{0}$, then $\mathbf{A} \circ \mathbf{B} \succeq \mathbf{0}$. Also, we have

$$\det(\mathbf{A} \circ \mathbf{B}) \geq \det(\mathbf{A}) \det(\mathbf{B}) \ .$$

The Kronecker Product

The *Kronecker product* is a generalization of the outer product which maps two matrices of arbitrary size to a block matrix. Specifically, if $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is a $p \times q$ matrix, then the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is an $mp \times nq$ block matrix given by

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{bmatrix} [\mathbf{A}]_{1,1}\,\mathbf{B} & \cdots & [\mathbf{A}]_{1,n}\,\mathbf{B} \\ \vdots & \ddots & \vdots \\ [\mathbf{A}]_{m,1}\,\mathbf{B} & \cdots & [\mathbf{A}]_{m,n}\,\mathbf{B} \end{bmatrix}.$$

Basic Properties:

- $\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) + (\mathbf{A} \otimes \mathbf{C})$.
- $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = (\mathbf{A} \otimes \mathbf{C}) + (\mathbf{B} \otimes \mathbf{C})$.
- $(\alpha\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (\alpha\mathbf{B}) = \alpha\,(\mathbf{A} \otimes \mathbf{B})$.
- $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$.
- $(\mathbf{A} \otimes \mathbf{B})^* = \mathbf{A}^* \otimes \mathbf{B}^*$, $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$, and $(\mathbf{A} \otimes \mathbf{B})^\dagger = \mathbf{A}^\dagger \otimes \mathbf{B}^\dagger$.
- There exist *permutation matrices* $\mathbf{P} \in \mathbb{F}^{mp \times mp}$, $\mathbf{Q} \in \mathbb{F}^{nq \times nq}$ such that

$$(\mathbf{A} \otimes \mathbf{B}) = \mathbf{P}\,(\mathbf{B} \otimes \mathbf{A})\,\mathbf{Q}\,.$$

Advanced Properties of the Kronecker Product

- **Products & Inverses:**
  - (*Mixed-Product Property:*) $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$.
  - By the mixed-product property, it follows that $(\mathbf{A} \otimes \mathbf{B})$ is invertible if and only if $\mathbf{A}$ and $\mathbf{B}$ are invertible, in which case we have

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} .$$

- **Spectrum:**
  Suppose $\mathbf{A} \in \mathbb{F}^{m \times m}$ and $\mathbf{B} \in \mathbb{F}^{n \times n}$. If $\lambda_1, \ldots, \lambda_m$ and $\mu_1, \ldots, \mu_n$ denote the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$, respectively, then the mn eigenvalues of $(\mathbf{A} \otimes \mathbf{B})$ are given by

$$\lambda_k \mu_\ell \, , \; 1 \leq k \leq m \, , \; 1 \leq \ell \leq n \, .$$

  From this, it follows that

$$\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A}) \cdot \mathrm{tr}(\mathbf{B}) \; , \; \det(\mathbf{A} \otimes \mathbf{B}) = (\det(\mathbf{A}))^n \, (\det(\mathbf{B}))^m \; .$$

- **Singular Values:**
  Suppose $\mathbf{A} \in \mathbb{F}^{m \times n}$ and $\mathbf{B} \in \mathbb{F}^{p \times q}$ have $\rho_{\mathbf{A}}$ and $\rho_{\mathbf{B}}$ nonzero singular values given by $\sigma_{\mathbf{A},1}, \ldots, \sigma_{\mathbf{A},\rho_{\mathbf{A}}}$ and $\sigma_{\mathbf{B},1}, \ldots, \sigma_{\mathbf{B},\rho_{\mathbf{B}}}$, respectively. Then, $(\mathbf{A} \otimes \mathbf{B})$ has $\rho_{\mathbf{A}} \rho_{\mathbf{B}}$ nonzero singular values given by

$$\sigma_{\mathbf{A},k} \sigma_{\mathbf{B},\ell} \, , \; 1 \leq k \leq \rho_{\mathbf{A}} \, , \; 1 \leq \ell \leq \rho_{\mathbf{B}} \, .$$

  From this it follows that

$$\mathrm{rank}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{rank}(\mathbf{A}) \cdot \mathrm{rank}(\mathbf{B}) \; .$$

Vectorization and the Vec Operator

In many cases, it is often convenient to stack the columns of a matrix together to form a column vector. This process of converting a matrix into a vector is called *vectorization* and is carried out using the *vec* operator. Specifically, if $\mathbf{A}$ is some $m \times n$ matrix with a column representation of the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \,,$$

where $\mathbf{a}_k$ is an $m \times 1$ vector for all k, then $\mathrm{vec}(\mathbf{A})$ is an $mn \times 1$ vector given by

$$\mathrm{vec}(\mathbf{A}) \triangleq \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$$

Equivalently, we have

$$[\mathrm{vec}(\mathbf{A})]_k = [\mathbf{A}]_{((k-1) \bmod m)+1, \left\lfloor \frac{k-1}{m} \right\rfloor + 1} \,, \; 1 \leq k \leq mn \,.$$

Properties of the Vec Operator

- **Compatibility with Kronecker products:**
  If $\mathbf{A} \in \mathbb{F}^{m \times n}$, $\mathbf{B} \in \mathbb{F}^{n \times p}$, and $\mathbf{C} \in \mathbb{F}^{p \times q}$, then we have

$$\begin{aligned} \mathrm{vec}(\mathbf{ABC}) &= \left(\mathbf{C}^T \otimes \mathbf{A}\right) \mathrm{vec}(\mathbf{B}) \,, \\ &= (\mathbf{I}_q \otimes \mathbf{AB}) \, \mathrm{vec}(\mathbf{C}) = \left(\mathbf{C}^T \mathbf{B}^T \otimes \mathbf{I}_m\right) \mathrm{vec}(\mathbf{A}) \,. \end{aligned}$$

- **Compatibility with Hadamard products:**
  If $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m \times n}$, then we have

$$\mathrm{vec}(\mathbf{A} \circ \mathbf{B}) = \mathrm{vec}(\mathbf{A}) \circ \mathrm{vec}(\mathbf{B}) \,.$$

- **Compatibility with inner products:** If $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m \times n}$, then we have

$$\mathrm{tr}\left(\mathbf{A}^\dagger \mathbf{B}\right) = (\mathrm{vec}(\mathbf{A}))^\dagger \, (\mathrm{vec}(\mathbf{B})) \,.$$

  In other words, for the standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^\dagger \mathbf{x}$ for vectors and $\langle \mathbf{X}, \mathbf{Y} \rangle = \mathrm{tr}\left(\mathbf{Y}^\dagger \mathbf{X}\right)$ for matrices, we have

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \langle \mathrm{vec}(\mathbf{X}) \,, \mathrm{vec}(\mathbf{Y}) \rangle \,.$$

Solving Lyapunov Equations Using the Vec Operator

In control theory, the stability of certain linear systems can be ascertained by solving a type of equation known as a *Lyapunov equation*. This equation takes one of the following forms:

$$\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^{\dagger} + \mathbf{Q} = \mathbf{0}\,, \quad \textbf{(continuous-time Lyapunov equation)}$$
$$\mathbf{A}\mathbf{X}\mathbf{A}^{\dagger} - \mathbf{X} + \mathbf{Q} = \mathbf{0}\,. \quad \textbf{(discrete-time Lyapunov equation)}$$

Here, all matrices are $n \times n$ and the stability for either the continuous or discrete-time systems can be determined by the value of $\mathbf{X}$. Specifically, in either case, if there exist $\mathbf{X} \succ \mathbf{0}$ and $\mathbf{Q} \succ \mathbf{0}$ such that the corresponding Lyapunov equation holds, then the linear system evolving according to the state matrix $\mathbf{A}$ is globally asymptotically stable.

The Lyapunov equation can be solved in closed form using the vec operator. Specifically, we have

$$\mathrm{vec}(\mathbf{X}) = -\left((\mathbf{I}_n \otimes \mathbf{A}) + (\mathbf{A}^* \otimes \mathbf{I}_n)\right)^{-1} \mathrm{vec}(\mathbf{Q}) \;, \quad \textbf{(continuous-time case)}$$
$$\mathrm{vec}(\mathbf{X}) = \left(\mathbf{I}_{n^2} - (\mathbf{A}^* \otimes \mathbf{A})\right)^{-1} \mathrm{vec}(\mathbf{Q}) \;. \quad \textbf{(discrete-time case)}$$

The Schur Complement: Formal Definition

Recall that if $\mathbf{A} \in \mathbb{F}^{m \times m}$, $\mathbf{B} \in \mathbb{F}^{m \times n}$, $\mathbf{C} \in \mathbb{F}^{n \times m}$, and $\mathbf{D} \in \mathbb{F}^{n \times n}$ are such that $\mathbf{A}$ and $\mathbf{D}$ are invertible, and if we define the $(m \times n) \times (m + n)$ matrix $\mathbf{M}$ in block form as follows:

$$\mathbf{M} \triangleq \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

then the $n \times n$ matrix $\mathbf{S}_{\mathbf{A};\mathbf{M}}$ defined as

$$\mathbf{S}_{\mathbf{A};\mathbf{M}} \triangleq \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B},$$

is said to be the *Schur complement* of $\mathbf{A}$ in $\mathbf{M}$. Similarly, the $m \times m$ matrix $\mathbf{S}_{\mathbf{D};\mathbf{M}}$ defined as $\mathbf{S}_{\mathbf{D};\mathbf{M}} \triangleq \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$ is the Schur complement of $\mathbf{D}$ in $\mathbf{M}$.

Schur complements arise in the study of block matrix inversion (i.e., when deriving an expression for $\mathbf{M}^{-1}$), and play a role in the *matrix inversion lemma*.

The Hermitian Case

Suppose that $\mathbf{A} \in \mathbb{H}^m$, $\mathbf{B} \in \mathbb{C}^{m \times n}$, and $\mathbf{C} \in \mathbb{H}^n$ are such that $\mathbf{A}$ is invertible. If we define the block matrix $\mathbf{X}$ as

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\dagger & \mathbf{C} \end{bmatrix},$$

and define the $n \times n$ matrix $\mathbf{S}$ as

$$\mathbf{S} \triangleq \mathbf{C} - \mathbf{B}^\dagger \mathbf{A}^{-1} \mathbf{B},$$

then we have the following results:

- $\mathbf{X} \succ \mathbf{0}$ if and only if $\mathbf{A} \succ \mathbf{0}$ and $\mathbf{S} \succ \mathbf{0}$.
- If $\mathbf{A} \succ \mathbf{0}$, then $\mathbf{X} \succeq \mathbf{0}$ if and only if $\mathbf{S} \succeq \mathbf{0}$.

These results can be easily shown by using the following decomposition:

$$\underbrace{\begin{bmatrix} \mathbf{v}_1^\dagger & \mathbf{v}_2^\dagger \end{bmatrix}}_{\mathbf{v}^\dagger} \underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\dagger & \mathbf{C} \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}}_{\mathbf{v}} =$$

$$\underbrace{\underbrace{\begin{bmatrix} \mathbf{v}_1^\dagger & \mathbf{v}_2^\dagger \end{bmatrix}}_{\mathbf{v}^\dagger} \overbrace{\begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m \times n} \\ \mathbf{B}^\dagger \mathbf{A}^{-1} & \mathbf{I}_n \end{bmatrix}}^{\mathbf{T}^\dagger}}_{\mathbf{u}^\dagger} \underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{0}_{m \times n} \\ \mathbf{0}_{n \times m} & \mathbf{S} \end{bmatrix}}_{\mathbf{Y}} \underbrace{\overbrace{\begin{bmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0}_{n \times m} & \mathbf{I}_n \end{bmatrix}}^{\mathbf{T}} \underbrace{\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}}_{\mathbf{v}}}_{\mathbf{u}}.$$

Case of Singular Block Term

If the matrix $\mathbf{A}$ from above is singular, then in general, it is not simple to relate the definiteness of $\mathbf{X}$ to its blocks, as the Schur complement of $\mathbf{A}$ in $\mathbf{X}$ is not formally defined in this case. However, if $\mathcal{R}(\mathbf{B}) \subseteq \mathcal{R}(\mathbf{A})$, then the following decomposition holds, based on a singular version of the Schur complement:

$$\mathbf{X} = \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m \times n} \\ \mathbf{B}^{\dagger}\mathbf{A}^{\#} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{0}_{m \times n} \\ \mathbf{0}_{n \times m} & \mathbf{C} - \mathbf{B}^{\dagger}\mathbf{A}^{\#}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{A}^{\#}\mathbf{B} \\ \mathbf{0}_{n \times m} & \mathbf{I}_n \end{bmatrix} .$$

It can be shown that

$$\mathcal{R}(\mathbf{B}) \subseteq \mathcal{R}(\mathbf{A}) \Longleftrightarrow \left(\mathbf{I}_m - \mathbf{A}\mathbf{A}^{\#}\right)\mathbf{B} = \mathbf{0} .$$

Hence, we have the following result:

- $\mathbf{X} \succeq \mathbf{0}$ if and only if $\mathbf{A} \succeq \mathbf{0}$, $\left(\mathbf{I}_m - \mathbf{A}\mathbf{A}^{\#}\right)\mathbf{B} = \mathbf{0}$, and $\mathbf{C} - \mathbf{B}^{\dagger}\mathbf{A}^{\#}\mathbf{B} \succeq \mathbf{0}$.

Expressing Constraints as Linear Matrix Inequalities

Often, constraints that one may encounter when formulating an optimization problem can be expressed in terms of linear matrix inequalities (LMIs). This can be extremely useful for posing the problem in terms of a *semidefinite program* (SDP), if possible, which is a standard form convex optimization problem.

Examples:

- Suppose $\mathbf{A} \in \mathbb{F}^{m \times n}$, $\mathbf{b} \in \mathbb{F}^{m \times 1}$, $\mathbf{x} \in \mathbb{F}^{n \times 1}$, and $\gamma > 0$, and we would like to enforce the vector norm constraint

  $$||\mathbf{A}\mathbf{x} - \mathbf{b}||_2 \leq \gamma \Longleftrightarrow \gamma^2 - (\mathbf{A}\mathbf{x} - \mathbf{b})^\dagger (\mathbf{A}\mathbf{x} - \mathbf{b}) \geq 0 \,.$$

  This can be expressed as an LMI as shown below:

  $$\begin{bmatrix} \mathbf{I}_m & (\mathbf{A}\mathbf{x} - \mathbf{b}) \\ (\mathbf{A}\mathbf{x} - \mathbf{b})^\dagger & \gamma^2 \end{bmatrix} \succeq \mathbf{0} \,.$$

- Suppose $\mathbf{X} \in \mathbb{F}^{m \times n}$ and $\gamma > 0$, and we would like to enforce the matrix norm constraint

  $$||\mathbf{X}||_2 = \sqrt{\lambda_{\max}\left(\mathbf{X}^\dagger \mathbf{X}\right)} \leq \gamma \,.$$

  As we have $\lambda_{\max}\left(\mathbf{X}^\dagger \mathbf{X}\right) \leq \gamma^2$ if and only if $\mathbf{X}^\dagger \mathbf{X} \preceq \gamma^2 \mathbf{I}_n$, this leads to the LMI given below:

  $$\begin{bmatrix} \mathbf{I}_m & \mathbf{X} \\ \mathbf{X}^\dagger & \gamma^2 \mathbf{I}_n \end{bmatrix} \succeq \mathbf{0} \,.$$

Algebraic Riccati Equations

Examples of equations which arises frequently in the study of control theory are *algebraic Riccati equations* (AREs), which determine the solution to the infinite horizon, time-invariant *linear-quadratic regulator* (LQR) and *linear-quadratic-Gaussian* (LQG) control problems. These equations come in the following variants:

$$\mathbf{A}^{\dagger}\mathbf{X} + \mathbf{X}\mathbf{A} - \mathbf{X}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^{\dagger}\mathbf{X} + \mathbf{Q} = \mathbf{0}\,, \qquad \textbf{(continuous-time ARE (CARE))}$$

$$\mathbf{X} = \mathbf{A}^{\dagger}\mathbf{X}\mathbf{A} + \mathbf{Q} - \left(\mathbf{A}^{\dagger}\mathbf{X}\mathbf{B}\right)\left(\mathbf{R} + \mathbf{B}^{\dagger}\mathbf{X}\mathbf{B}\right)^{-1}\left(\mathbf{B}^{\dagger}\mathbf{X}\mathbf{A}\right). \qquad \textbf{(discrete-time ARE (DARE))}$$

Here, $\mathbf{A} \in \mathbb{F}^{m \times m}$, $\mathbf{B} \in \mathbb{F}^{m \times n}$, $\mathbf{Q} \in \mathbb{H}^{m}$, and $\mathbf{R} \in \mathbb{H}^{n}$ are the problem data and $\mathbf{X} \in \mathbb{H}^{m}$ is the variable to be solved. Typically we assume $\mathbf{R} \succ \mathbf{0}$ and we seek a solution for which $\mathbf{X} \succeq \mathbf{0}$. To find such a solution, it is common to *relax* an equality to an inequality. For the DARE, one such relaxation is given below:

$$\mathbf{X} \preceq \mathbf{A}^{\dagger}\mathbf{X}\mathbf{A} + \mathbf{Q} - \left(\mathbf{A}^{\dagger}\mathbf{X}\mathbf{B}\right)\left(\mathbf{R} + \mathbf{B}^{\dagger}\mathbf{X}\mathbf{B}\right)^{-1}\left(\mathbf{B}^{\dagger}\mathbf{X}\mathbf{A}\right)\,,\ \mathbf{X} \succeq \mathbf{0}\,.$$

This can be shown to lead to the following LMI:

$$\begin{bmatrix} \mathbf{R} + \mathbf{B}^{\dagger}\mathbf{X}\mathbf{B} & \mathbf{B}^{\dagger}\mathbf{X}\mathbf{A} & \mathbf{0}_{n \times m} \\ \mathbf{A}^{\dagger}\mathbf{X}\mathbf{B} & \mathbf{A}^{\dagger}\mathbf{X}\mathbf{A} + \mathbf{Q} - \mathbf{X} & \mathbf{0}_{m \times m} \\ \mathbf{0}_{m \times n} & \mathbf{0}_{m \times m} & \mathbf{X} \end{bmatrix} \succeq \mathbf{0}\,.$$