Compound Matrices and Dynamical Systems

Michael Margaliot

Joint work with:

Eyal Bar-Shalom, Omri Dalin, Rami Katz, Ron Ofir,
A. Ovseevich, F. Bullo, E. D. Sontag.

Disclaimer: I will forsake rigor for intuition.

Questions most welcome any time.

Lessons from Gian-Carlo Rota

1. Every lecture should make only one main point.
2. Never run overtime.
3. Relate to your audience.
4. Give them something to take home.

Main point

Compound matrices play a significant role in systems and control theory.

k-minors of a matrix

Let $A \in \mathbb{R}^{n\times m}$. Fix $k \in \{1, \dots, \min\{n, m\}\}$. Fix indices:

$$1 \le i_1 < i_2 < \cdots < i_k \le n,$$
$$1 \le j_1 < j_2 < \cdots < j_k \le m.$$

The corresponding *k-submatrix* of A is

$$B = \begin{bmatrix} A(i_1, j_1) & A(i_1, j_2) & \dots & A(i_1, j_k) \\ A(i_2, j_1) & A(i_2, j_2) & \dots & A(i_2, j_k) \\ & & \vdots & \\ A(i_k, j_1) & A(i_k, j_2) & \dots & A(i_k, j_k) \end{bmatrix}.$$

A corresponding *k-minor* is

$$\det(B).$$

k-minors of a matrix

Let $A \in \mathbb{R}^{n \times m}$. Fix $k \in \{1, \dots, \min\{n, m\}\}$. Fix indices:

$$1 \le i_1 < i_2 < \dots < i_k \le n,$$
$$1 \le j_1 < j_2 < \dots < j_k \le m.$$

$$A = \begin{bmatrix} A(1,1) & A(1,2) & \dots & A(1,m) \\ A(2,1) & A(2,2) & \dots & A(2,m) \\ & & \vdots & \\ A(n,1) & A(n,2) & \dots & A(n,m) \end{bmatrix}.$$

The number of k-minors is

$$\binom{n}{k} \times \binom{m}{k}.$$

k-multiplicative compound of a matrix

Definition

The k *multiplicative compound* of a matrix $A \in \mathbb{R}^{n \times m}$ is a matrix $A^{(k)} \in \mathbb{R}^{\binom{n}{k} \times \binom{m}{k}}$ that collects all the k-minors of A (in a lexicographic ordering).

Note $A^{(1)} = A$. If $A \in \mathbb{R}^{n \times n}$ then $A^{(n)} = \det(A)$.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \implies A^{(2)} = \begin{bmatrix} \det(\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}) & \det(\begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}) & \det(\begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}) \\ & \vdots & \\ \det(\begin{bmatrix} 4 & 7 \\ 5 & 8 \end{bmatrix}) & \det(\begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}) & \det(\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}) \end{bmatrix}.$$

k-multiplicative compound of a diagonal matrix

Definition

The k *multiplicative compound* of a matrix $A \in \mathbb{R}^{n\times m}$ is a matrix $A^{(k)} \in \mathbb{R}^{\binom{n}{k}\times\binom{m}{k}}$ that collects all the k-minors of A (in a lexicographic ordering).

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \implies A^{(2)} = \begin{bmatrix} \det(\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}) & \det(\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}) & \det(\begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}) \\ & \vdots & \\ \det(\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}) & \det(\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}) & \det(\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}) \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 12 \end{bmatrix}.$$

Eigenvalues of A: 2, 3, and 4. Eigenvalues of $A^{(2)}$: $2 * 3$, $2 * 4$, and $3 * 4$.

Properties of the k-multiplicative matrix

Theorem (Cauchy-Binet)

Let $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{m\times p}$. Fix $k \in \{1, \dots, \min\{n, m, p\}\}$. Then

$$(AB)^{(k)} = A^{(k)} B^{(k)}.$$

For $k = 1$, this becomes $AB = AB$.

For $A, B \in \mathbb{R}^{n\times n}$ and $k = n$, this becomes $\det(AB) = \det(A)\det(B)$.

The Cauchy-Binet theorem justifies the terminology "multiplicative compound".

Inverse of the k-multiplicative matrix

Theorem (Cauchy-Binet)

Let $A, B \in \mathbb{R}^{n \times n}$. Fix $k \in \{1, \dots, n\}$. Then

$$(AB)^{(k)} = A^{(k)} B^{(k)}.$$

Let I_p denote the $p \times p$ identity matrix. Suppose that $A \in \mathbb{R}^{n \times n}$ is non-singular. Then $I_n = AA^{-1}$, so

$$(I_n)^{(k)} = (AA^{-1})^{(k)} = A^{(k)}(A^{-1})^{(k)},$$

that is,

$$I_{\binom{n}{k}} = A^{(k)}(A^{-1})^{(k)},$$

so $(A^{(k)})^{-1} = (A^{-1})^{(k)}$.

Spectral properties of the k-multiplicative compound

Theorem

Denote the eigenvalues of $A \in \mathbb{R}^{n\times n}$ by: $\lambda_1, \dots, \lambda_n$. Then the eigenvalues of $A^{(k)}$ are the $\binom{n}{k}$ products:

$$\lambda_{i_1}\lambda_{i_2}\dots\lambda_{i_k}, \quad \text{with } 1 \le i_1 < i_2 < \dots < i_k \le n.$$

For $k = n$, this gives $A^{(n)} = \det(A) = \lambda_1 \dots \lambda_n$.

Proof.

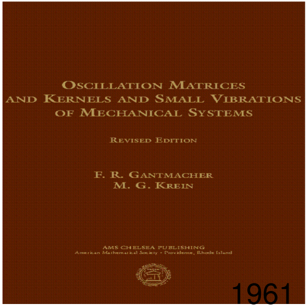
Suppose that $D := TAT^{-1}$ is diagonal. Then

$$D^{(k)} = (TAT^{-1})^{(k)} = T^{(k)}A^{(k)}(T^{-1})^{(k)} = T^{(k)}A^{(k)}(T^{(k)})^{-1}.$$

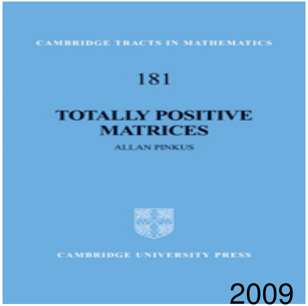
Applications of multiplicative compounds

Definition

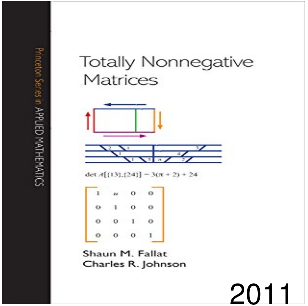
A matrix $A \in \mathbb{R}^{n\times n}$ is called *totally positive* (TP) if all its minors are positive.



1961



2009



2011

Applications of multiplicative compounds

Definition

A matrix $A \in \mathbb{R}^{n \times n}$ is called *totally positive* (TP) if all its minors are positive.

For example, consider $A = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix}$.
Its 1-minors (i.e., entries) are positive.
It has a single 2-minor: $\det(A) = 2$, so it is TP.

Theorem (Gantmacher and Krein)

If $A \in \mathbb{R}^{n \times n}$ is TP then all its eigenvalues are real, positive, and simple.

Applications of multiplicative compounds

Theorem (Gantmacher and Krein)

If $A \in \mathbb{R}^{n\times n}$ is TP then all its eigenvalues are real, positive, and simple.

Proof.

Since 1-minors are positive, the Perron Theorem implies that there exists a positive eigenvalue λ_1 such that

$$\lambda_1 > |\lambda_j| \text{ for all } j \neq 1. \tag{1}$$

Since $A^{(2)}$ is positive,

$$\lambda_1\lambda_2 > |\lambda_j\lambda_\ell| \text{ for all } (j,\ell) \neq (1,2).$$

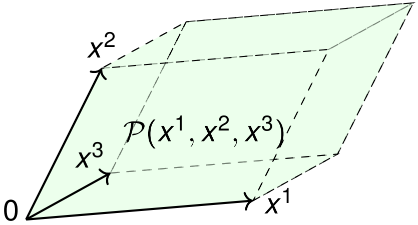
so λ_2 is real and positive, and $\lambda_1 > \lambda_2 > |\lambda_3|$. Continue to $A^{(3)}$, ... ∎

Applications of multiplicative compounds

Definition

Fix $x^1, \dots, x^k \in \mathbb{R}^n$. The *parallelotope generated by* $x^1, \dots, x^k$ (and the origin) is

$$P(x^1, \dots, x^k) := \{\sum_{i=1}^{k} r_i x^i : r_i \in [0, 1]\}.$$



Applications of multiplicative compounds

Definition

Fix $x^1, \dots, x^k \in \mathbb{R}^n$. The *parallelotope generated by* $x^1, \dots, x^k$ (and the origin) is

$$P(x^1, \dots, x^k) := \{\sum_{i=1}^{k} r_i x^i : r_i \in [0,1]\}.$$

Theorem (Gantmacher, 1960)

Let $A := \begin{bmatrix} x^1 & \dots & x^k \end{bmatrix} \in \mathbb{R}^{n \times k}$. *Then the volume of* $P(x^1, \dots, x^k)$ *is equal to*

$$| \begin{bmatrix} x^1 & \dots & x^k \end{bmatrix}^{(k)} |_2.$$

Note that $\begin{bmatrix} x^1 & \dots & x^k \end{bmatrix}^{(k)} \in \mathbb{R}^{\binom{n}{k} \times \binom{k}{k}}$.

Applications of multiplicative compounds

Theorem (Gantmacher, 1960)

Let $A := \begin{bmatrix} x^1 & \dots & x^k \end{bmatrix} \in \mathbb{R}^{n \times k}$. *Then the volume of* $P(x^1, \dots, x^k)$ *is equal to*

$$|\begin{bmatrix} x^1 & \dots & x^k \end{bmatrix}^{(k)}|_2.$$

For $k = n$, this gives

$$\text{volume}(P(x^1, \dots, x^n)) = |\begin{bmatrix} x^1 & \dots & x^n \end{bmatrix}^{(n)}|_2 = |\det(\begin{bmatrix} x^1 & \dots & x^n \end{bmatrix})|.$$

Compound matrices in Linear Ordinary Differential Equations

Consider the ODE:

$$\dot{x}(t) = A(t)x(t), \quad x(0) = x_0.$$

The solution is

$$x(t, x_0) = \Phi(t)x_0,$$

where $\Phi(t)$ is the transition matrix from time 0 to time t. The transition matrix is the solution of the matrix linear ODE:

$$\dot{\Phi}(t) = A(t)\Phi(t), \quad \Phi(0) = I_n.$$

Question: what can we say about $(\Phi(t))^{(k)}$?

Compound matrices in Linear ODEs

Consider the ODE:

$$\dot{x}(t) = A(t)x(t), \quad x(0) = x_0.$$

Fix k initial conditions $a^1, \dots, a^k \in \mathbb{R}^n$. Then

$$\begin{bmatrix} x(t, a^1) & \dots & x(t, a^k) \end{bmatrix}^{(k)}$$

describes the evolution in time of the parallelotope generated by $a^1, \dots, a^k$. Note that

$$\begin{aligned} \begin{bmatrix} x(t, a^1) & \dots & x(t, a^k) \end{bmatrix}^{(k)} &= \begin{bmatrix} \Phi(t)a^1 & \dots & \Phi(t)a^k \end{bmatrix}^{(k)} \\ &= \begin{bmatrix} \Phi(t) \begin{bmatrix} a^1 & \dots & a^k \end{bmatrix} \end{bmatrix}^{(k)} \\ &= (\Phi(t))^{(k)} \begin{bmatrix} a^1 & \dots & a^k \end{bmatrix}^{(k)}. \end{aligned}$$

Thus, $(\Phi(t))^{(k)}$ describes the evolution of the k-parallelotopes under the ODE.

Compound matrices in Linear ODEs

The transition matrix satisfies $\dot{\Phi}(t) = A(t)\Phi(t)$, with $\Phi(0) = I_n$. We would like to study the evolution of $(\Phi(t))^{(k)}$. What is $\frac{d}{dt}(\Phi(t))^{(k)}$?

$$
\begin{aligned}
(\Phi(t+\varepsilon))^{(k)} - (\Phi(t))^{(k)} &\approx (\Phi(t) + \varepsilon\dot{\Phi}(t))^{(k)} - (\Phi(t))^{(k)} \\
&= (\Phi(t) + \varepsilon A(t)\Phi(t))^{(k)} - (\Phi(t))^{(k)} \\
&= ((I_n + \varepsilon A(t))\Phi(t))^{(k)} - (\Phi(t))^{(k)} \\
&= (I_n + \varepsilon A(t))^{(k)}(\Phi(t))^{(k)} - (\Phi(t))^{(k)} \\
&= \left((I_n + \varepsilon A(t))^{(k)} - I_{\binom{n}{k}}\right)(\Phi(t))^{(k)}.
\end{aligned}
$$

Define the k additive compound of $A \in \mathbb{R}^{n\times n}$ by

$$
A^{[k]} := \lim_{\varepsilon \to 0} \frac{(I_n + \varepsilon A)^{(k)} - I_{\binom{n}{k}}}{\varepsilon}.
$$

Compound matrices in Linear ODEs

The transition matrix satisfies

$$\dot{\Phi}(t) = A(t)\Phi(t), \quad \Phi(0) = I_n,$$

and

$$\frac{d}{dt}(\Phi(t))^{(k)} = (A(t))^{[k]}(\Phi(t))^{(k)}, \quad (\Phi(0))^{(k)} = I_{\binom{n}{k}}.$$

Thus, the evolution of k-dimensional parallelotopes also follows a linear ODE with the k additive compound matrix $(A(t))^{[k]}$.
For $k = n$, this becomes

$$\frac{d}{dt}\det(\Phi(t)) = \operatorname{trace}(A(t))\det(\Phi(t)).$$

Properties of the k additive compound

Theorem

Denote the eigenvalues of $A \in \mathbb{R}^{n \times n}$ by: $\lambda_1, \dots, \lambda_n$. Then the eigenvalues of $A^{[k]}$ are the $\binom{n}{k}$ sums:

$$\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k}, \quad \text{with } 1 \leq i_1 < i_2 < \cdots < i_k \leq n.$$

For $k = n$, this gives $A^{[n]} = \lambda_1 + \cdots + \lambda_n = \text{trace}(A)$.

Theorem

Let $A, B \in \mathbb{R}^{n \times n}$. Then $(A + B)^{[k]} = A^{[k]} + B^{[k]}$.

This justifies the terminology additive compound.

Applications to nonlinear systems

Consider the nonlinear system:

$$\dot{x}(t) = f(x(t)). \tag{2}$$

The first step is to derive a *linear system* that provides useful information on (2). Fix two initial conditions $a, b \in \mathbb{R}^n$, and let $x(t,a)$, $x(t,b)$ denote the corresponding solutions at time t. Let $z(t) := x(t,a) - x(t,b)$. Then

$$\dot{z}(t) = \left(\int_0^1 \frac{\partial}{\partial x} f(rx(t,a) + (1-r)x(t,b))\, \mathrm{d}r \right) z(t).$$

This *variational equation* is a linear ODE. Let $J := \frac{\partial}{\partial x} f$.
Then $J^{[k]}$ describes the evolution of k-dimensional parallelotopes under the variational equation.

k generalizations of nonlinear dynamical systems using k compounds

Consider the nonlinear system:

$$\dot{x}(t) = f(x(t)).$$

Let $J(x) := \frac{\partial}{\partial x} f(x)$.

A generalization principle

Suppose that the system satisfies some property $\iff$ $J(x)$ satisfies a condition p.
We say that the system satisfies k*-property* $\iff$ $(J(x))^{[k]}$ satisfies condition p.

This is meaningful because:

1. For $k = 1$, we have $J^{[k]} = J$, so we obtain the original property,
2. $J^{[k]}$ has a clear geometric interpretation.

From contraction to k-contraction

The system $\dot{x} = f(x)$ is called *contractive* if for any two initial conditions a, b the corresponding solutions approach each other at an exponential rate:

$$|x(t,a) - x(t,b)| \leq \exp(-\eta t)|a - b|, \text{ for all } t \geq 0,$$

where $\eta > 0$.

Implications:

- If the system admits an equilibrium e then e is globally exponentially stable,
- If the system admits more than a single equilibrium point then it is not contractive,
- If $\dot{x} = f(t, x)$ and f is T-periodic then the system admits a unique globally exponentially stable T-periodic solution $\gamma(t)$, $t \in [0, T)$.

From contraction to k contraction

Consider the nonlinear system $\dot{x} = f(x)$.
There is a simple sufficient condition for contraction. For a norm $|\cdot| : \mathbb{R}^n \to \mathbb{R}_+$ and a matrix $A \in \mathbb{R}^{n \times n}$, define the induced matrix norm

$$\|A\| := \max_{|x|=1} |Ax|,$$

and the induced matrix measure

$$\mu(A) := \lim_{\varepsilon \downarrow 0} \frac{\|I_n + \varepsilon A\| - 1}{\varepsilon}.$$

Theorem

If $\mu(J(x)) \leq -\eta < 0$ for all x then the nonlinear system is contractive.

Contraction theory has found numerous applications in robotics, multi-agent systems, dynamic neural network models, systems biology, and more.

From contraction to k contraction

Consider the nonlinear system $\dot{x} = f(x)$.

$\mu(J(x)) \leq -\eta < 0$ – the system is contractive
$\mu((J(x))^{[k]}) \leq -\eta < 0$ – the system is k-contractive

k-contraction implies that along the variational equation k-dimensional parallelotopes converge to zero at an exponential rate.

From contraction to 2-contraction

Theorem (Muldowney and Li)

If the time-invariant system $\dot{x}(t) = f(x(t))$ is 2-contractive then any bounded solution converges to an equilibrium point.

This allows to prove a well-behaved asymptotic behaviour in nonlinear systems that admit more than a single equilibrium, e.g., dynamical neural network models that serve as associative memories.

Consider

$$\dot{x} = A(t)x,$$

with

$$A(t) = \begin{bmatrix} -1 & 0 \\ -2\cos(t) & 0 \end{bmatrix} x.$$

For any $x(0) \in \mathbb{R}^2$, we have

$$\lim_{t \to \infty} x(t, x(0)) = \begin{bmatrix} 0 \\ x_2(0) - x_1(0) \end{bmatrix}.$$

Thus, there is more than a single equilibrium point, so the system is not contractive. However, $A^{[2]}(t) = \text{trace}(A(t)) \equiv -1$, so the system is 2-contractive.

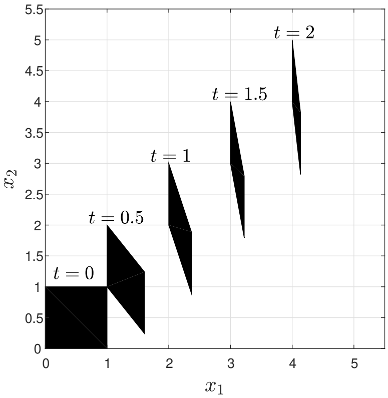


Figure: Evolution of the unit square in a 2-contractive system.

2-contraction in networked systems

Consider the nonlinear networked system

$$\dot{x}(t) = -\begin{bmatrix} d_1 & 0 & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & d_n \end{bmatrix} x(t) + W_1 f\left(W_2 x(t)\right) + v, \tag{3}$$

with $v \in \mathbb{R}^n$, $W_1, W_2 \in \mathbb{R}^{n \times n}$. Let $J_f(x) := \frac{\partial}{\partial x} f(x)$. Fix $k \in \{1, \dots, n\}$, and let

$$\alpha_k := k^{-1} \min \left\{ d_{i_1} + \dots + d_{i_k} \mid 1 \le i_1 < \dots < i_k \le n \right\}.$$

Theorem (Ofir, Ovseevich, Margaliot)

If

$$\alpha_k > 0 \text{ and } \sup_x \| J_f(W_2 x) \|_2^2 \sum_{i=1}^{k} \sigma_i^2(W_1) \sigma_i^2(W_2) < \alpha_k^2 k$$

then (3) *is k-contractive.*

Contractive System $\mu(J) \le -\eta < 0$	k-Contractive System $\mu(J^{[k]}) \le -\eta < 0$
Positive System A is Metzler	k-Positive System $A^{[k]}$ is Metzler
Cooperative System J is Metzler	k-Cooperative System $J^{[k]}$ is Metzler
Discrete Time Diagonal Stable A is diagonally stable	Discrete Time k-Diagonal Stable $A^{(k)}$ is diagonally stable
Hankel 1-Positive System $G(z)$ is externally positive	Hankel k-Positive System $G^{(k)}(z)$ is externally positive

Main point

Compound matrices play a significant role in systems and control theory.

References

- E. Bar-Shalom, O. Dalin, and M. Margaliot, "Compound matrices in systems and control theory: a tutorial," *Mathematics of Control, Signals, and Systems*, 2023.
- M. Margaliot and E. D. Sontag„ "Revisiting totally positive differential systems: A tutorial and new results," *Automatica*, 2019.
- C. Wu, I. Kanevskiy, and M. Margaliot, "k-contraction: theory and applications," *Automatica*, 2022.
- J. S. Muldowney, "Compound matrices and ordinary differential equations," *Rocky Mountain J. Math.*, 1990.
- R. Ofir, A. Ovseevich, and M. Margaliot, "Contraction and k-contraction in Lurie systems with applications to networked systems," *Automatica*, 2024.
- C. Grussler, and R. Sepulchre, "Variation diminishing linear time-invariant systems," *Automatica*, 2022.