Matrix Rank-One Decomposition

and Applications

Shuzhong Zhang

Department of Systems Engineering and Engineering Management
The Chinese University of Hong Kong

ICNPA, Beijing
April 8, 2008

Outline

- Trust Region Subproblems
- Radar Code Selection
- The Matrix Rank-One Decomposition
- Theoretical Applications

Trust Region Subproblem

The Trust-Region Subproblem:

$$\begin{array}{ll} \text{minimize} & x^{\mathrm{T}} Q_0 x - 2 b_0^{\mathrm{T}} x \\ \text{subject to} & \|x\| \le \delta. \end{array}$$

The CDT Trust Region Subproblem

The CDT (Celis, Dennis, Tapia, 1985) Trust-Region Subproblem:

$$\begin{array}{ll} \text{minimize} & x^{\mathrm{T}} Q_0 x - 2 b_0^{\mathrm{T}} x \\ \text{subject to} & \|Ax - b\| \le \delta_1 \\ & \|x\| \le \delta_2. \end{array}$$

The Radar Code Selection Problem

(Based on De Maio, De Nicola, Huang, Z., Farina, 2007)

A radar system transmits a coherent burst of pulses

$$s(t) = a_t u(t) \exp\left(\boldsymbol{i}(2\pi f_0 t + \phi)\right)$$

- a_t is the transmit signal amplitude;
- $u(t) = \sum_{k=0}^{N-1} a(k)p(t - kT_r)$ is the signal's complex envelope;
- $p(t)$ is the signature of the transmitted pulse, and T_r is the Pulse Repetition Time (PRT);
- $[a(0), a(1), \ldots, a(N-1)] \in C^N$ is the radar code (assumed without loss of generality with unit norm);
- f_0 is the carrier frequency, and ϕ is a random phase.

The Output

The filter output is

$$v(t) = \alpha_r e^{-\boldsymbol{i}2\pi f_0 \tau} \sum_{k=0}^{N-1} a(k) e^{\boldsymbol{i}2\pi k f_d T_r} \chi_p(t - kT_r - \tau, f_d) + w(t)$$

where $\chi_p(\lambda, f)$ is the pulse waveform ambiguity function

$$\chi_p(\lambda, f) = \int_{-\infty}^{+\infty} p(\beta) p^*(\beta - \lambda) e^{\boldsymbol{i}2\pi f \beta} d\beta$$

and $w(t)$ is the down-converted and filtered disturbance component.

Sampling

The signal $v(t)$ is sampled at $t_k = \tau + kT_r$, $k = 0, \ldots, N-1$, the output becomes

$$v(t_k) = \alpha a(k) e^{\boldsymbol{i} 2\pi k f_d T_r} \chi_p(0, f_d) + w(t_k), \qquad k = 0, \ldots, N-1$$

where $\alpha = \alpha_r e^{-\boldsymbol{i} 2\pi f_0 \tau}$.

Denote

$$\boldsymbol{c} = [a(0), a(1), \ldots, a(N-1)]^{\mathrm{T}},$$

$$\boldsymbol{p} = [1, e^{\boldsymbol{i} 2\pi f_d T_r}, \ldots, e^{\boldsymbol{i} 2\pi (N-1) f_d T_r}]^{\mathrm{T}} \text{ (the temporal steering vector)}$$

$$\boldsymbol{w} = [w(t_0), w(t_1), \ldots, w(t_{N-1})]^{\mathrm{T}}$$

the backscattered signal can be written as

$$\boldsymbol{v} = \alpha \boldsymbol{c} \odot \boldsymbol{p} + \boldsymbol{w}$$

where $\odot$ denotes the Hadamard product.

Performance, Doppler Accuracy, and Similarity

The Optimal Code Design Problem can be formulated as

$$\begin{cases} \max_{\boldsymbol{c}} & \boldsymbol{c}^{\mathrm{H}} \boldsymbol{R} \boldsymbol{c} \\ \text{s.t.} & \boldsymbol{c}^{\mathrm{H}} \boldsymbol{c} = 1 \\ & \boldsymbol{c}^{\mathrm{H}} \boldsymbol{R}_1 \boldsymbol{c} \geq \delta_a \\ & \|\boldsymbol{c} - \boldsymbol{c}_0\|^2 \leq \epsilon \end{cases}$$

where $\boldsymbol{R} = \Gamma^{-1} \odot (\boldsymbol{p}^{\mathrm{H}} \boldsymbol{p})$ with $\Gamma = \mathsf{E}[\boldsymbol{w}\boldsymbol{w}^{\mathrm{H}}]$, and $\boldsymbol{R}_1 = \Gamma^{-1} \odot (\boldsymbol{p}\boldsymbol{p}^{\mathrm{H}})^* \odot (\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}})^*$ with $\boldsymbol{u} = [0, \boldsymbol{i}2\pi, \ldots, \boldsymbol{i}2\pi(N-1)]^{\mathrm{T}}$.

Commonalities

Non-Convex Quadratically Constrained Quadratic Optimization (QCQP), in real and/or complex variables, with a few constraints.

Matrix Rank-One Decomposition

Theorem (Sturm and Z.; 2003).
Let $A \in \mathcal{S}^n$. Let $X \in \mathcal{S}^n_+$ with rank r. There exists a rank-one decomposition for X such that

$$X = \sum_{i=1}^{r} x_i x_i^{\mathrm{T}}$$

and $x_i^{\mathrm{T}} A x_i = \frac{A \bullet X}{r}$, $i = 1, ..., r$.

Can we do more?

It is easy to show by example that in general it is only possible to get a complete rank-one decomposition with respect to one matrix. But it is possible to get a partial decomposition for two:

Theorem (Ai and Z.; 2006).
Let $A_1, A_2 \in \mathcal{S}^n$ and $X \in \mathcal{S}^n_+$. If $r := \text{rank}(X) \geq 3$ then one can find in polynomial-time (real-number sense) a rank-one decomposition for X,

$$X = x_1 x_1^{\mathrm{T}} + x_2 x_2^{\mathrm{T}} + \cdots + x_r x_r^{\mathrm{T}},$$

such that

$$A_1 \bullet x_i x_i^{\mathrm{T}} = \frac{A_1 \bullet X}{r}, \quad i = 1, ..., r$$

$$A_2 \bullet x_i x_i^{\mathrm{T}} = \frac{A_2 \bullet X}{r}, \quad i = 1, ..., r-2.$$

The Hermitian case

Theorem (Huang and Z.; 2006).
Let $A_1, A_2 \in \mathcal{H}^n$, and $X \in \mathcal{H}^n_+$ with rank r. There exists a rank-one decomposition for X such that

$$X = \sum_{i=1}^{r} x_i x_i^{\mathrm{H}}$$

and $x_i^{\mathrm{H}} A_k x_i = \frac{A_k \bullet X}{r}$, $i = 1, ..., r$; $k = 1, 2$.

Analog in the Hermitian case

Theorem (Ai, Huang and Z.; 2007).
Suppose that $A_1, A_2, A_3 \in \mathcal{H}^n$ and $X \in \mathcal{H}^n_+$. If $r = \mathrm{rank}(X) \geq 3$, then one can find in polynomial-time (real-number sense) a rank-one decomposition for X,

$$X = \sum_{i=1}^{r} x_i x_i^{\mathrm{H}},$$

such that

$$A_1 \bullet x_i x_i^{\mathrm{H}} = \delta_1 / r, A_2 \bullet x_i x_i^{\mathrm{H}} = \delta_2 / r, \text{ for all } i = 1, \ldots, r;$$
$$A_3 \bullet x_i x_i^{\mathrm{H}} = \delta_3 / r, \text{ for } i = 1, \ldots, r-2.$$

Solving QP by Matrix Decomposition

Quadratically Constrained Quadratic Programming (QCQP):

$$
\begin{array}{lll}
(Q) & \text{minimize} & q_0(x) = x^{\mathrm{H}} Q_0 x - 2\mathrm{Re}\ b_0^{\mathrm{H}} x \\
 & \text{subject to} & q_i(x) = x^{\mathrm{H}} Q_i x - 2\mathrm{Re}\ b_i^{\mathrm{H}} x + c_i \le 0, \quad i = 1, ..., m.
\end{array}
$$

SDP Relaxation

Let

$$M(q_0) := \begin{bmatrix} 0 & -b_0^{\mathrm{H}} \\ -b_0 & Q_0 \end{bmatrix}, \; M(q_i) := \begin{bmatrix} c_i & -b_i^{\mathrm{H}} \\ -b_i & Q_i \end{bmatrix}, \text{ for } i = 1, ..., m.$$

Then, (Q) is equivalently written as

$$\begin{array}{lll} (Q) & \min & M(q_0) \bullet \begin{bmatrix} t \\ x \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}^{\mathrm{H}} = x^{\mathrm{H}} Q_0 x - 2\mathrm{Re}\; b_0^{\mathrm{H}} x \bar{t} \\ & \text{s.t.} & M(q_i) \bullet \begin{bmatrix} t \\ x \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}^{\mathrm{H}} = x^{\mathrm{H}} Q_i x - 2 b_i^{\mathrm{H}} x \bar{t} + c_i |t|^2 \le 0, \quad i = 1, ..., m \\ & & |t|^2 = 1. \end{array}$$

SDP Relaxation

The so-called SDP relaxation of (Q) is

$$\begin{array}{lll} (SP) & \text{minimize} & M(q_0) \bullet X \\ & \text{subject to} & M(q_i) \bullet X \leq 0, \quad i = 1, ..., m \\ & & I_{00} \bullet X = 1 \\ & & X \succeq 0 \quad \boxed{X \text{ rank one}} \end{array}$$

where $I_{00} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{0} \end{bmatrix} \in \mathcal{H}^{n+1}$. The dual problem of (SP) is:

$$\begin{array}{lll} (SD) & \text{maximize} & y_0 \\ & \text{subject to} & Z = M(q_0) - y_0 I_{00} + \sum_{i=1}^m y_i M(q_i) \succeq 0 \\ & & y_i \geq 0, i = 1, ..., m. \end{array}$$

Complementary Slackness

Under suitable conditions, (SP) and (SD) have complementary optimal solutions, X^* and Z^*:

$$X^* Z^* = \mathbf{0}.$$

If we can decompose X^* into *rank-one* summations, evenly satisfying all the constraints, then *each of the rank-one vectors will be optimal!*

Consequences of the Matrix Decomposition Theorems

Polynomially solvable cases of the nonconvex quadratic programs:

Real quadratic program:

$m = 1$ ($m = 2$ if homogeneous) $\Longleftarrow$ (Sturm & Z., 2003)

Real quadratic program:

$m = 2$ ($m = 3$ if h.) $\text{rank}(X^*) \geq 3$ $\Longleftarrow$ (Ai & Z., 2006)

Complex quadratic program:

$m = 2$ ($m = 3$ if h.) $\Longleftarrow$ (Huang & Z., 2005)

Complex quadratic program:

$m = 3$ ($m = 4$ if h.) $\text{rank}(X^*) \geq 3$ $\Longleftarrow$ (Ai, Huang & Z., 2007)

Further Theoretical Applications

Field of Values of a Matrix

Let A be any $n \times n$ matrix, the *field of values* of A is given by

$$\mathcal{F}(A) := \{z^{\mathrm{H}} A z \mid z^{\mathrm{H}} z = 1\} \subseteq \mathbf{C}.$$

This set, like the spectrum set, contains a lot of information about the matrix A.

The set is known to be convex.

Reference: R.A. Horn and C.R. Johnson. *Topics in Matrix analysis.* Cambridge University Press, Cambridge, 1991.

Joint Numerical Ranges

In general, the *joint numerical range* of matrices is defined to be

$$\mathcal{F}(A_1, ..., A_m) := \left\{ \left(\begin{array}{c} z^{\mathrm{H}} A_1 z \\ \vdots \\ z^{\mathrm{H}} A_m z \end{array} \right) \middle| \; z^{\mathrm{H}} z = 1, z \in \mathbf{C}^n \right\}.$$

Theorem (Hausdorff; 1919).
If A_1 and A_2 are Hermitian, then $\mathcal{F}(A_1, A_2)$ is a convex set.

A Theorem of Brickman

Theorem (Brickman; 1961).

Suppose that A_1, A_2, A_3 are $n \times n$ Hermitian matrices. Then

$$\left\{ \left(\begin{array}{c} z^{\mathrm{H}} A_1 z \\ z^{\mathrm{H}} A_2 z \\ z^{\mathrm{H}} A_3 z \end{array} \right) \middle| \; z \in \mathbf{C}^n \right\}$$

is a convex set.

The S-Procedure

It is often useful to consider the following implication

$$G_1(x) \geq 0, G_2(x) \geq 0, ..., G_m(x) \geq 0 \Longrightarrow F(x) \geq 0.$$

A sufficient condition is:

$$\exists \tau_1 \geq 0, \tau_2 \geq 0, ..., \tau_m \geq 0 \text{ such that } F(x) - \sum_{i=1}^{m} \tau_i G_i(x) \geq 0 \, \forall x.$$

This procedure is called *lossless* if the above condition is also *necessary.*

The S-Lemma

Theorem (Jakubovic; 1971).
Suppose that $m = 1$, and F, G_1 are real quadratic forms. Moreover, there is $x_0 \in \Re^n$ such that $x_0^{\mathrm{T}} G_1 x_0 > 0$. Then the S-procedure is lossless.

Theorem (Jakubovic; 1971).
Suppose that $m = 2$, and F, G_1, G_2 are Hermitian quadratic forms. Moreover, there is $x_0 \in \mathbf{C}^n$ such that $x_0^{\mathrm{H}} G_i x_0 > 0$, $i = 1, 2$. Then the S-procedure is lossless.

Proof of the S-Lemma: The Hermitian case

We need only to show that the S-procedure is lossless in this case. Let $G_i(x) = x^{\mathrm{H}} A_i x$, $i = 1, 2$, and $F(x) = x^{\mathrm{H}} A_3 x$.

Consider the following cone

$$\left\{ \left(\begin{array}{c} x^{\mathrm{H}} A_1 x \\ x^{\mathrm{H}} A_2 x \\ x^{\mathrm{H}} A_3 x \end{array} \right) \middle| \, x \in \mathbf{C}^n \right\}.$$

It is a convex cone in $\Re^3$ by Brickman's theorem.

Moreover, it does not intersect with $\Re_{++} \times \Re_{++} \times \Re_{--}$.

Proof of the S-Lemma (continued)

By the separation theorem, there is $(t_1, t_2, t_3) \neq 0$, such that

$$t_1 x_1 + t_2 x_2 + t_3 x_3 \leq 0, \, \forall x_1 > 0, x_2 > 0, x_3 < 0,$$

and

$$t_1 x^{\mathrm{H}} A_1 x + t_2 x^{\mathrm{H}} A_2 x + t_3 x^{\mathrm{H}} A_3 x \geq 0, \, \forall x \in \mathbf{C}^n.$$

The first condition implies that $t_1 \leq 0$, $t_2 \leq 0$, and $t_3 \geq 0$. We see that $t_3 > 0$ in this case, and so

$$A_3 - \frac{t_1}{t_3} A_1 - \frac{t_2}{t_3} A_2 \succeq 0.$$

But how to prove Brickman's theorem?

Clearly, it will be sufficient if we can show

$$\left\{ \left(\begin{array}{c} z^{\mathrm{H}} A_1 z \\ z^{\mathrm{H}} A_2 z \\ z^{\mathrm{H}} A_3 z \end{array} \right) \middle| \; z \in \mathbf{C}^n \right\} = \left\{ \left(\begin{array}{c} A_1 \bullet Z \\ A_2 \bullet Z \\ A_3 \bullet Z \end{array} \right) \middle| \; Z \succeq 0 \right\}$$

Proof of the Brickman Theorem

Take any nonzero vector

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} A_1 \bullet Z \\ A_2 \bullet Z \\ A_3 \bullet Z \end{pmatrix}.$$

Suppose that $v_3 \neq 0$. Consider two matrix equations

$$\left(A_1 - \frac{v_1}{v_3} A_3\right) \bullet Z = 0$$

$$\left(A_2 - \frac{v_2}{v_3} A_3\right) \bullet Z = 0$$

Proof of the Brickman Theorem (continued)

Using our decomposition, there will be $Z = \sum_{i=1}^{r} z_i z_i^{\mathrm{H}}$ such that

$$z_i^{\mathrm{H}} \left(A_1 - \tfrac{v_1}{v_3} A_3 \right) z_i = 0$$

$$z_i^{\mathrm{H}} \left(A_2 - \tfrac{v_2}{v_3} A_3 \right) z_i = 0$$

for $i = 1, ..., r$. Among these, there will be one vector such that $z_i^{\mathrm{H}} A_3 z_i$ has the same sign as $A_3 \bullet Z$.

Let $\rho := \sqrt{v_3 / z_i^{\mathrm{H}} A_3 z_i}$, and $z := \rho z_i$. Then,

$$z^{\mathrm{H}} A_3 z = \rho^2 z_i^{\mathrm{H}} A_3 z_i = v_3, \; z^{\mathrm{H}} A_k z = \frac{v_k}{v_3} z^{\mathrm{H}} A_3 z = v_k, \; k = 1, 2.$$

A Result of Yuan

Theorem (Yuan; 1990).

Let A_1 and A_2 be in $\mathcal{S}^n$. If

$$\max\{x^{\mathrm{T}} A_1 x, x^{\mathrm{T}} A_2 x\} \ge 0 \; \forall x \in \Re^n$$

then there exist $\mu_1 \ge 0, \mu_2 \ge 0, \mu_1 + \mu_2 = 1$ such that

$$\mu_1 A_1 + \mu_2 A_2 \succeq 0.$$

Extension

Theorem (Ai, Huang, Zhang; 2007).
Let A_1, A_2, A_3 be in $\mathcal{H}^n$. If

$$\max\{z^{\mathrm{H}} A_1 z, z^{\mathrm{T}} A_2 z, z^{\mathrm{T}} A_3 z\} \geq 0 \; \forall z \in C^n$$

then there exist $\mu_1, \mu_2, \mu_3 \geq 0, \mu_1 + \mu_2 + \mu_3 = 1$ such that

$$\mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 \succeq 0.$$

Further Extension

Theorem (Ai, Huang, Zhang; 2007).
Suppose that $n \geq 3$, $A_i \in \mathcal{H}^n, i = 1,2,3,4$, and there are $\lambda_i \in \Re, i = 1,2,3,4$, such that $\lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3 + \lambda_4 A_4 \succ 0$. If

$$\max\{z^{\mathrm{H}} A_1 z, z^{\mathrm{H}} A_2 z, z^{\mathrm{H}} A_3 z, z^{\mathrm{H}} A_4 z\} \geq 0, \; \forall z \in \mathbf{C}^n$$

then there are $\mu_i \geq 0, i = 1,2,3,4$, such that $\mu_1 + \mu_2 + \mu_3 + \mu_4 = 1$ such that

$$\mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 + \mu_4 A_4 \succeq 0.$$