# THE MATRIX CUBE PROBLEM: Approximations and Applications

Arkadi Nemirovski,
Stieltjes Visiting Professor, TU Delft
and
Technion – Israel Institute of Technology
nemirovs@ie.technion.ac.il

joint research with A. Ben-Tal

1. Matrix Cube
   - The problem: formulation and motivation
   - Main result
   - Back to applications
   - Sketch of the proof
2. From Matrix Cube to Computing Matrix Norms
   - The problem
   - Main result
   - Sketch of the proof

## The Matrix Cube Problem

♣ **In several applications, we meet with the** *Matrix Cube* **problems:**

♠ **MatrCube.A:** *Given $n \times n$ symmetric matrices $B^0, B^1, ..., B^L$ and $\rho \geq 0$, check whether the "matrix box"*

$$\mathcal{U}[\rho] = \left\{ B = B^0 + \sum_{\ell=1}^{L} u_\ell B^\ell : |u_\ell| \leq \rho, \, \ell = 1, ..., L \right\}$$

*is $\succeq 0$, i.e., all matrices from $\mathcal{U}[\rho]$ are positive semidefinite.*

♠ **MatrCube.B:** *Find the largest $\rho \geq 0$ such that $\mathcal{U}[\rho] \succeq 0$.*

**Applications:**

- **Numerous problems of Robust Control, e.g., Lyapunov Stability Analysis for uncertain dynamical systems;**
- **Various combinatorial problems which can be reduced to maximizing a positive definite quadratic form over the unit cube.**

## Lyapunov Stability Analysis

♠ **Consider an uncertain linear time-varying dynamical system**

$$\dot{x}(t) = A(t)x(t), \quad t \geq 0. \qquad \textbf{(S)}$$

**Here** $A(t)$ **is not known exactly. All we know is that the entries in** $A(t)$ **belong to given "uncertainty intervals":**

$$A(t) \in \mathcal{A}_\rho \equiv \left\{ A : |A_{ij} - A^\star_{ij}| \leq \rho D_{ij},\, 1 \leq i, j \leq n \right\}.$$

<u>**Question:**</u> **How to certify that (S) is** *stable* **– all trajectories of the system converge to 0 as** $t \to \infty$ **?**

<u>**Answer:**</u> **Try to find a** *quadratic Lyapunov stability certificate* **– a matrix** $X \succeq I$ **such that**

$$A^T X + XA \preceq -I \quad \forall (A \in \mathcal{A}_\rho).$$

$$\boxed{\begin{array}{c}\exists X: \\ X \succeq I \ \&\ A^T X + XA \preceq -I \quad \forall (A \in \mathcal{A}_\rho) \ \textbf{(L)} \\ \Downarrow \\ \textbf{System } \dot{x}(t) = A(t)x(t),\ A(\cdot) \in \mathcal{A}[\rho]\textbf{, is stable}\end{array}}$$

$\diamondsuit$ **If $X$ satisfies the premise, then**

$$\exists \alpha > 0: \ A^T X + XA \preceq -\alpha X \quad \forall A \in \mathcal{A}[\rho]$$

$$\begin{array}{rcl}\Rightarrow \frac{d}{dt}(z^T(t)Xz(t)) & = & z^T(t)[A^T(t)X + XA(t)]z(t) \\ & \leq & -\alpha z^T(t)Xz(t)\end{array}$$

$$\Rightarrow z^T(t)Xz(t) \leq e^{-\alpha t} z^T(0)Xz(0)$$

$$\Rightarrow z(t) \to 0,\ t \to \infty.$$

♣ **To find efficiently an $X$ satisfying (L) is, essentially, the same as to be able to check whether** *a given* $X$ **satisfies (L). This is nothing but problem MatrCube.A:**

$$\boxed{\begin{array}{l}\forall (u : |u_{ij}| \leq \rho) : \\ [A^\star_{ij} + u_{ij}D_{ij}]^T X + X[A^\star_{ij} + u_{ij}D_{ij}] \preceq -I\end{array}}$$

$$\Updownarrow$$

$$\boxed{\begin{array}{l}\forall (u : |u_{ij}| \leq \rho) : \\ \underbrace{[-I - (A^\star)^T X - XA^\star]}_{B^0 = B^0[X]} \\ \qquad\qquad + \sum\limits_{i,j} u_{ij} \underbrace{D_{ij}[e_j(Xe_i)^T + (Xe_i)e_j^T]}_{B^{ij} = B^{ij}[X]} \succeq 0\end{array}}$$

## Maximizing Positive Definite Quadratic Form over Unit Cube

♠ **Let $S \succ 0$. Consider the problem**

$$\omega(S) = \max_x \left\{ x^T S x : \|x\|_\infty \leq 1 \right\}. \qquad \textbf{(Q)}$$

**Lemma:** *Let* $\mathcal{U}[\rho] = \left\{ A = A^T : |A_{ij} - (S^{-1})_{ij}| \leq \rho \right\}$**.** *Then*

$$\omega^{-1}(S) = \max\left\{ \rho : \mathcal{U}[\rho] \succeq 0 \right\}.$$

*Thus,* (Q) *is a very specific particular case of* **MatrCube.B.**

**Proof:**

$$\begin{array}{l} \omega(S) = \min\left\{ \omega : \|x\|_S^2 \equiv x^T S x \leq \omega \|x\|_\infty^2 \quad \forall x \right\} \\ = \min\left\{ \omega : \|\xi\|_{S^{-1}}^2 \equiv \xi^T S^{-1} \xi \geq \omega^{-1} \|\xi\|_1^2 \quad \forall \xi \right\} \\ = \min\left\{ \omega : \xi^T S^{-1} \xi \geq \omega^{-1} \max\limits_{B = B^T : |B_{ij}| \leq 1} \xi^T B \xi \quad \forall \xi \right\} \\ = \min\left\{ \omega : S^{-1} - \omega^{-1} B \succeq 0 \quad \forall (B : |B_{ij}| \leq 1) \right\} \\ = \frac{1}{\max\{\rho : \mathcal{U}[\rho] \succeq 0\}}. \end{array}$$

## Intermediate Summary

♠ <u>Good news:</u> **MatrCube has important applications.**

♠ <u>Bad news:</u> **MatrCube is NP-hard (since (Q) is so).**

♠ <u>Good news:</u> **Although NP-hard, problem MatrCube admits simple tractable approximation.**

**<u>Lemma:</u>** *Let* $\mathcal{U}[\rho] = \left\{ B^0 + \sum_{\ell=1}^{L} u_\ell B^\ell : |u_\ell| \le \rho \right\}$. *Assume that the system of LMIs*

$$\begin{array}{rcl} X_\ell & \succeq & \pm B^\ell,\ \ell = 1, ..., L, \\ \rho \sum_{\ell=1}^{L} X_\ell & \preceq & B^0 \end{array} \qquad (\mathrm{A}[\rho])$$

*in matrix variables* $X_1, ..., X_L$ *is solvable. Then* $\mathcal{U}[\rho] \succeq 0$, *i.e., the answer in* **MatrCube.A** *is affirmative.*

**<u>Corollary:</u>** *The efficiently computable quantity*

$$\widehat{\rho} = \max \{ \rho : \ (\mathrm{A}[\rho]) \textit{ is solvable} \}$$

*is a lower bound on the optimal value* $\rho^\star$ *in problem* **MatrCube.B**.

**Proof of Lemma: If $X_\ell$ are such that**

$$\begin{array}{ll}(a) & X_\ell \succeq \pm B^\ell,\ \ell=1,...,L,\\ (b) & \rho\sum_{\ell=1}^{L} X_\ell \preceq B^0\end{array}$$

**and**

$$B = B^0 + \sum_{\ell=1}^{L} u_\ell B^\ell,\ |u_\ell| \le \rho$$

**is a matrix from $\mathcal{U}_\rho$, then**

$$\begin{array}{rcll} B &=& B^0 + \sum_{\ell=1}^{L} u_\ell B^\ell & \\ &\succeq& B^0 - \sum_{\ell=1}^{L} \rho X_\ell & [\textbf{by } (a) \textbf{ due to } |u_\ell| \le \rho] \\ &\succeq& 0 & [\textbf{by } (b)] \end{array}$$

**Thus, $\mathcal{U}[\rho] \succeq 0$.**

$$\boxed{\exists\{X_\ell\}:\begin{cases} X_\ell \succeq \pm B^\ell \\ \rho\sum_\ell X_\ell \preceq B^0 \end{cases} \qquad (\mathrm{II}[\rho])}$$

$$\Downarrow$$

$$\boxed{\mathcal{U}[\rho] \equiv \left\{B^0 + \sum_\ell u_\ell B^\ell : |u_\ell| \le \rho\right\} \succeq 0 \quad (\mathrm{I}[\rho])}$$

♡ **<u>Matrix Cube Theorem:</u>** *Efficiently verifiable sufficient condition* (II[$\rho$]) *for "intractable" predicate* (I[$\rho$]) *is tight, provided that the ranks of the "edge matrices"* $B^1, ..., B^L$ *are small. Specifically, if*

$$\mu = \max_{1\le\ell\le L} \mathrm{Rank}(B^\ell)$$

(note $\ell \ge 1$ in max!) *and* $\rho \ge 0$ *is such that* (II[$\rho$]) *does <u>not</u> take place, then so is* (I[$\vartheta(\mu)\rho$]). *Here* $\vartheta(\mu)$ *is an universal function such that*

$$\vartheta(1) = 1, \vartheta(2) = \frac{\pi}{2} \approx 1.57, \vartheta(3) \approx 1.73, \vartheta(4) = 2$$

*and*

$$\vartheta(\mu) \le \frac{\pi\sqrt{\mu}}{2} \quad \forall \mu.$$

*In particular,*

$$1 \le \frac{\rho^\star}{\widehat{\rho}} = \frac{\max\{\rho : (\mathrm{I}[\rho]) \text{ is valid}\}}{\max\{\rho : (\mathrm{II}[\rho]) \text{ is valid}\}} \le \vartheta(\mu).$$

**♣ In the Matrix Cube problems responsible for Interval Lyapunov Stability Analysis and for Quadratic maximization over the unit cube, the ranks of the "edge matrices" are at most 2. In light of the Matrix Cube Theorem, it follows that**

♢ *One can efficiently bound from below the largest level $\rho^\star$ of uncertainty for which all instances of an interval matrix*

$$\mathcal{A}_\rho = \{A : |A_{ij} - A^\star_{ij}| \leq \rho D_{ij}\}$$

*share a common quadratic Lyapunov stability certificate. The bound is*

$$\widehat{\rho} = \max_{\rho, X, X^{ij}} \left\{ \rho : \begin{array}{r} X^{ij} \succeq \pm D_{ij}[(Xe_i)e_j^T + e_j(Xe_i)^T] \\ i,j = 1,...,n, \\ \rho \sum\limits_{i,j} X^{ij} \preceq -I - (A^\star)^T X - XA^\star \\ X \succeq I \end{array} \right\}$$

*and is tight within the factor $\vartheta(2) = \frac{\pi}{2}$.*

**Similar results are valid for other problems of Robust Control under interval uncertainty:**

- **Lyapunov stability synthesis,**
- **Robust dissipativity analysis,**
- **Synthesis of robust optimal controllers in Linear-Quadratic Control, etc.**

$\diamondsuit$ *One can efficiently bound from above the maximum $\omega(S)$ of a positive definite quadratic form $x^T Sx$ over the unit cube. The bound is given by*

$$\widehat{\omega}^{-1} = \max\left\{\rho : \begin{array}{r} (1+\delta_{ij})X^{ij} \succeq \pm[e_i e_j^T + e_j e_i^T], \\ 1 \le i \le j \le n, \\ \rho \sum\limits_{1\le i\le j\le n} X_{ij} \preceq S^{-1} \end{array}\right\}$$

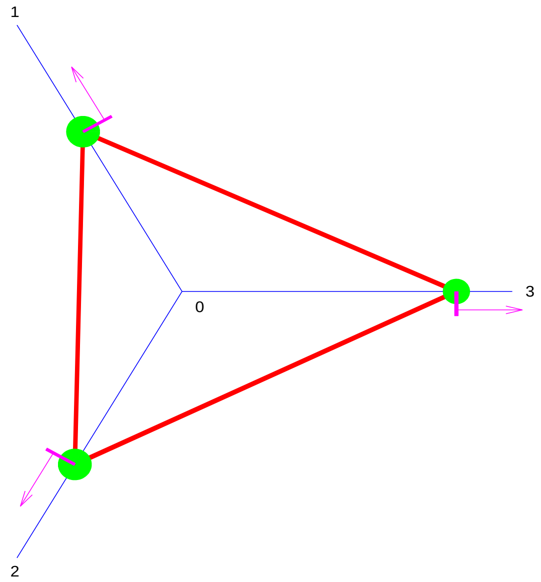
*and is tight within the factor $\vartheta(2) = \frac{\pi}{2}$.*

**On a closest inspection, $\widehat{\omega}$ turns out to be the well-known** *semidefinite relaxation* **bound:**

$$\begin{array}{rcl} \widehat{\omega} & = & \max\limits_X \{\mathrm{Tr}(SX) : X \succeq 0,\ X_{ii} \le 1\} \\ & = & \min\limits_\lambda \left\{\sum\limits_i \lambda_i : S \preceq \mathrm{Diag}\{\lambda\}\right\}. \end{array}$$

**The fact that this bound is tight within the factor $\frac{\pi}{2}$ was originally established by Yu. Nesterov (1997) via completely different approach originating from the MAXCUT-related "random hyperplane" technique of Goemans and Williamson (1995).**

♠ **Example:** *Three material points, linked by elastic springs, can slide with friction along three axes in 2D plane:*

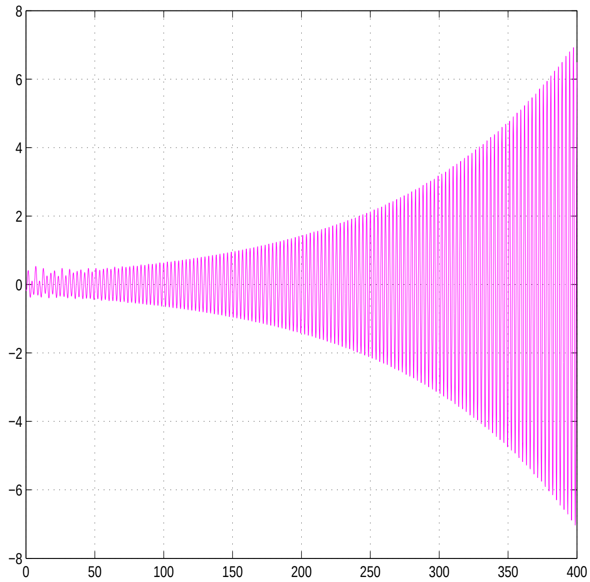



*Given the nominal masses of the points, rigidities of the springs, friction coefficients and equilibrium positions of the points, what is the largest level $\rho^{\rm safe}$ of stability-preserving time-varying perturbations of masses and rigidities?*

♡ **With the outlined approach, it turns out that at the level of perturbations $\widehat{\rho} = 0.38\%$ all perturbed instances of the system share a common quadratic Lyapunov stability certificate. Thus,**

$$\rho^{\rm safe} \geq \widehat{\rho} = 0.0038.$$

♡ **Numerical experiments demonstrate that with time-varying 2.3% perturbations the system may lose stability:**




**Sample shift of point # 1 vs. time, perturbations 2.3%**

**Thus,**

$$\rho^{safe} \leq 0.023 = 6\widehat{\rho}.$$

## Sketch of the Proof

♣ **Situation:** We are given an integer $\mu$ and a real $\rho \geq 0$ such that the ranks of the matrices $B^1, ..., B^L$ do not exceed $\mu$ and the system of LMIs

$$\begin{array}{rcl} X_\ell & \succeq & \pm B^\ell \\ \rho \sum\limits_\ell X_\ell & \preceq & B^0 \end{array} \qquad (\mathrm{A}[\rho])$$

in matrix variables $X_\ell$ has no solution.

♣ **Target:** To prove the existence of $u_\ell$, $|u_\ell| \leq \vartheta(\mu)\rho$, such that the matrix

$$B = B^0 + \sum_{\ell=1}^{L} u_\ell B^\ell$$

is not positive semidefinite. Here $\vartheta(\cdot)$ is given by

$$\vartheta^{-1}(\mu) = \min_\alpha \left\{ \int\limits_{\mathbf{R}^\mu} \left| \sum_{i=1}^{\mu} \alpha_i \xi_i^2 \right| p_\mu(\xi) d\xi : \|\alpha\|_1 \geq 1 \right\},$$

where $p_\mu(\cdot)$ is the standard Gaussian density on $\mathbf{R}^\mu$ (zero mean, unit covariance matrix). It is quite straightforward to verify that this function possesses the properties, like $\vartheta(2) = \frac{\pi}{2}$, $\vartheta(\mu) \leq \frac{\pi\sqrt{\mu}}{2}$, etc., announced in the Matrix Cube Theorem.

♣ <u>Step 1 (routine):</u> **From the fact that the system of LMIs**

$$\begin{array}{rcl} X_\ell & \succeq & \pm B^\ell \\ \rho \sum\limits_\ell X_\ell & \preceq & B^0 \end{array} \qquad (\mathrm{A}[\rho])$$

**in matrix variables** $X_\ell$ **has no solution it follows by** *semidefinite duality* **that**

$$\begin{array}{l} \exists Y \succeq 0 : \\ \rho \sum\limits_{\ell=1}^m \|\lambda(\underbrace{Y^{1/2} B^\ell Y^{1/2}}_{C_\ell})\|_1 > \mathrm{Tr}(\underbrace{Y^{1/2} B^0 Y^{1/2}}_{C_0}); \end{array} \qquad (\star)$$

**here** $\lambda(C)$ **is the vector of eigenvalues of a symmetric matrix** $C$**.**

$$\exists Y \succeq 0 :$$
$$\rho \sum_{\ell=1}^{m} \|\lambda(\underbrace{Y^{1/2}B^{\ell}Y^{1/2}}_{C_\ell})\|_1 > \mathrm{Tr}(\underbrace{Y^{1/2}B^{0}Y^{1/2}}_{C_0}). \qquad (\star)$$

♣ **Step 2 (crucial):** *The quantities* $\|\lambda(C_\ell)\|_1$, $\ell = 1, ..., L$, *and* $\mathrm{Tr}(C_0)$ *admit probabilistic interpretation.* **Specifically, let** $\xi \sim \mathcal{N}(0, I_n)$**.**

♡ **If** $C$ **is an** $n \times n$ **symmetric matrix, then**
  **(a)** $\mathbf{E}\left\{|\xi^T C\xi|\right\} \geq \|\lambda(C)\|_1 \vartheta^{-1}(\mathrm{Rank}(C))$**;**
  **(b)** $\mathbf{E}\left\{\xi^T C\xi\right\} = \mathrm{Tr}(C)$**.**

**(b) is evident. Due to the rotational invariance of the standard Gaussian distribution, it suffices to verify (a) in the case when** $C$ **is** *diagonal***; in this case, (a) is readily given by the definition of** $\vartheta(\cdot)$**.**

♢ **By (a), (b) relation** $(\star)$ **implies that**

$$\mathbf{E}\left\{\rho\vartheta(\mu)\sum_{\ell=1}^{L}|\xi^T Y^{1/2}B^{\ell}Y^{1/2}\xi| - \xi^T Y^{1/2}B^{0}Y^{1/2}\xi\right\} > 0\,.$$

$\diamondsuit$ **We now have**

$$\mathbf{E}\left\{\rho\vartheta(\mu)\sum_{\ell=1}^{L}|\xi^TY^{1/2}B^\ell Y^{1/2}\xi| - \xi^TY^{1/2}B^0Y^{1/2}\xi\right\} > 0$$

$$\Downarrow$$

$$\exists\zeta: \quad \rho\vartheta(\mu)\sum_{\ell=1}^{L}|\zeta^TB^\ell\zeta| - \zeta^TB^0\zeta > 0$$

$$\Downarrow$$

$$\zeta^T[B^0 + \sum_{\ell=1}^{L}\underbrace{\rho\vartheta(\mu)\epsilon_\ell}_{u_\ell,|u_\ell|\leq\vartheta(\mu)\rho} B^\ell]\zeta < 0$$

$$\left[\epsilon_\ell = -\operatorname{sign}(\zeta^TB^\ell\zeta)\right]$$

$$\Downarrow$$

$$\mathcal{U}[\vartheta(\mu)\rho] \not\succeq 0.$$

## From Matrix Cube to Matrix Norms

♣ **In fact, problem MatrCube.B:**

$$\max\left\{\rho : B^0 + \sum_{\ell=1}^{L} u_\ell B^\ell \succeq 0 \quad \forall (u : \|u\|_\infty) \le \rho\right\}$$

**asks to compute a specific norm of a linear mapping. Indeed, w.l.o.g. we may assume that $B^0 \succ 0$. Rewriting the problem as**

$$\rho^\star = \max\left\{\rho : I + \sum_{\ell=1}^{L} u_\ell \bar{B}^\ell \succeq 0 \quad \forall (u : \|u\|_\infty \le \rho\right\}$$
$$\left[\bar{B}^\ell = (B^0)^{-1/2} B^\ell (B^0)^{-1/2}\right]$$

**we see that our task is** *to find the norm of the linear map*

$$u \mapsto B(u) = \sum_{\ell=1}^{L} u_\ell \bar{B}^\ell : \mathbf{R}^L \to \mathbf{S}^n$$

*when the argument space $\mathbf{R}^L$ is equipped with the norm $\|u\|_\infty = \max_\ell |u_\ell|$, and the image space $\mathbf{S}^n$ of symmetric $n \times n$ matrices is equipped with the standard matrix norm $\|A\| = \|\lambda(A)\|_\infty$*:

$$(\rho^\star)^{-1} = \max\left\{\|B(u)\| : \|u\|_\infty \le 1\right\}.$$

**Indeed,**

$$\begin{aligned}
\rho^\star &= \max\left\{\rho : I + B(u) \succeq 0 \ \forall (u : \|u\|_\infty \le \rho)\right\} \\
&= \max\left\{\rho : \begin{array}{c} I + B(-u) \succeq 0, \\ I + B(u) \succeq 0 \end{array} \ \forall (u : \|u\|_\infty \le \rho)\right\} \\
&= \max\left\{\rho : -I \preceq B(u) \preceq I \ \forall (u : \|u\|_\infty \le \rho)\right\} \\
&= \max\left\{\rho : \|B(u)\| \le 1 \ \forall (u : \|u\|_\infty \le \rho)\right\} \\
&= \frac{1}{\max\{\|B(u)\| : \|u\|_\infty \le 1\}}.
\end{aligned}$$

♣ **What about a seemingly simpler "Matrix Norm" problem as follows:**

> **MatrNorm**$(p,r)$**:** *Given an $m\times n$ real matrix $A$ and reals $p,r\in[1,\infty]$, compute the norm of the linear mapping*
>
> $$x\mapsto Ax:\mathbf{R}^n\to\mathbf{R}^m$$
>
> *when the argument space is equipped with the norm $\|\cdot\|_p$, and the image space – with the norm $\|\cdot\|_r$, i.e., find the quantity*
>
> $$\begin{array}{rcl}\|A\|_{p,r} & \equiv & \max\limits_{x\neq 0}\frac{\|Ax\|_r}{\|x\|_p}\\ & = & \max\limits_{x}\left\{\|Ax\|_r:\|x\|_p\leq 1\right\}\\ & = & \max\limits_{y,x}\left\{y^TAx:\|x\|_p\leq 1,\|y\|_{r_*}\leq 1\right\}\\ & & \left[s_*=\frac{s}{s-1}\Leftrightarrow\frac{1}{s}+\frac{1}{s_*}=1\right]\\ & = & \|A^T\|_{r_*,p_*}.\end{array}$$

**In this problem, $A$ is the data, and $p,r$ are once for ever fixed "structural parameters". W.l.o.g., we may assume that $A$ is square.**

♣ **The extremely simple-looking problem** $\textsf{MatrNorm}(p, r)$ **is not that simple. It asks to** *maximize the convex function* $\|Ax\|_r$ **over the convex set** $\{\|x\|_p \leq 1\}$**, which, in general, is a fairly difficult task. In fact, we know only 3 simple cases of the problem:**

♡ $\underline{p = 1}$**:** $\|A\|_{1,r} = \max\limits_{1\leq j\leq n} \|a_j\|_r,\ A = [a_1, ..., a_n]$.

**The formula merely says that the maximum of a convex function** $\|Ax\|_r$ **on the** *polyhedral* **set** $\{x : \|x\|_1 \leq 1\}$ **is attained at a vertex, and there are just** $2n$ **vertices, the** $\pm$ **basic orths.**

♡ $\underline{r = \infty}$**:** $\|A\|_{p,\infty} = \max\limits_{1\leq i\leq n} \|a_i\|_{p*}, \quad A = \begin{bmatrix} a_1^T \\ \ldots \\ a_n^T \end{bmatrix}$.

**This case is "symmetric" to the one of** $p = 1$**. Recall that**

$$\|A\|_{p,r} = \|A^T\|_{r*,p*},$$

**so that the "computability status" of** $\textsf{MatrNorm}(p, r)$ **is exactly the same as the one of** $\textsf{MatrNorm}(r_*, p_*)$**.**

♡ <u>$p = r = 2$</u>**:** $\|A\|_{2,2} = \sqrt{\lambda_{\max}(A^T A)}$.
**This is the self-symmetric case of the standard matrix norm.**

♠ <u>**Bad news:**</u> *Problem* **MatrNorm**$(p, r)$ *is NP-hard when* $r < p$.

♠ <u>**Conjecture:**</u> *The cases of* $p = 1$, *of* $r = \infty$ *and of* $p = r = 2$ *are the* <u>*only*</u> *cases when problem* **MatrNorm**$(p, r)$*is easy; in all remaining cases it is NP-hard.*

♣ <u>**Good news:**</u> *When* $p \geq 2 \geq r$, *problem* **MatrNorm**$(p, r)$, *although NP-hard (except for* $p = r = 2$*), admits reasonably tight computationally tractable approximations.*

## Semidefinite Relaxation of $\textsf{MatrNorm}(p,r)$, $p \geq 2 \geq r$

♣ **Let** $a_1^T, ...a_n^T$ **be the rows of an** $n\times n$ **matrix** $A$**, and let** $p \geq 2 \geq r$**. Setting**

$$X[x] = xx^T,$$

**and denoting** $d(X)$ **the diagonal of a square matrix** $X$**, we have**

$$\begin{array}{l}
\|A\|_{p,r} = \max\limits_x \{\|Ax\|_r : \|x\|_p \leq 1\} \\
= \max\limits_x \left\{ \left( \sum\limits_{i=1}^n [(a_i^T x)^2]^{\frac{r}{2}} \right)^{\frac{1}{r}} : \|d(X[x])\|_{\frac{p}{2}} \leq 1 \right\} \\
= \max\limits_x \left\{ \left( \sum\limits_{i=1}^n [a_i^T X[x] a_i]^{\frac{r}{2}} \right)^{\frac{1}{r}} : \|d(X[x])\|_{\frac{p}{2}} \leq 1 \right\} \\
= \max\limits_X \left\{ \left( \sum\limits_{i=1}^n [a_i^T X a_i]^{\frac{r}{2}} \right)^{\frac{1}{r}} : \begin{array}{l} \|d(X)\|_{\frac{p}{2}} \leq 1 \\ X \succeq 0 \\ \mathrm{Rank}(X) = 1 \end{array} \right\} \\
\leq \max\limits_X \left\{ \left( \sum\limits_{i=1}^n [a_i^T X a_i]^{\frac{r}{2}} \right)^{\frac{1}{r}} : \begin{array}{l} \|d(X)\|_{\frac{p}{2}} \leq 1 \\ X \succeq 0 \end{array} \right\}.
\end{array}$$

♡ <u>**Conclusion:**</u> *The quantity*

$$\Omega_{p,r}(A) = \max_X \left\{ \left( \sum_{i=1}^n [a_i^T X a_i]^{\frac{r}{2}} \right)^{\frac{1}{r}} : \begin{array}{l} \|d(X)\|_{\frac{p}{2}} \leq 1 \\ X \succeq 0 \end{array} \right\}$$

*is an upper bound on* $\|A\|_{p,r}$.

**Since** $p \geq 2$**, the function** $\|d(X)\|_{\frac{p}{2}}$ **is convex in** $X$**. Since** $r \leq 2$**, the function**

$$\left( \sum_{i=1}^n [a_i^T X a_i]^{\frac{r}{2}} \right)^{\frac{1}{r}}$$

**is concave in** $X \succeq 0$**. Thus,**

♡ *The bound* $\Omega_{p,r}(A)$ *is the optimal value in an explicit Convex Programming problem and as such is efficiently computable.*

♡ <u>**Nice fact:**</u> *When* $p \geq 2 \geq r$*, the bound* $\Omega_{p,r}$ *is intelligent enough to recognize the identity*

$$\|A\|_{p,r} = \|A^T\|_{r_*,p_*}.$$

*Specifically,*

$$\Omega_{p,r}(A) = \Omega_{r_*,p_*}(A^T).$$

## Tightness of the Bound $\Omega_{p,r}$

♡ **<u>Matrix Norm Theorem:</u>** *Let*

$$\infty \geq p \geq 2 \geq r \geq 1.$$

*Then for every $n \times n$ matrix $A$ one has*

$$\|A\|_{p,r} \leq \Omega_{p,r}(A) \leq \underbrace{\min\left[\frac{\Phi(p,n)}{\Psi(r_*)}, \frac{\Phi(r_*,n)}{\Psi(p)}\right]}_{\Theta(p,r,n)} \cdot \|A\|_{p,r},$$

*where for $w \geq 2$*

$$\begin{array}{rcl}\Phi(w,n) &=& \min\left[\sqrt{2}\left(\frac{\Gamma\left(\frac{w+1}{2}\right)}{\sqrt{\pi}}\right)^{1/w}, \sqrt{2\ln(n+1)}\right] \\ &\leq& \sqrt{\min\left[2w-1, 2\ln(n+1)\right]}, \\ \Psi(w) &=& \sqrt{2}\left(\frac{\Gamma\left(\frac{2w-1}{2w-2}\right)}{\sqrt{\pi}}\right)^{\frac{w-1}{w}} \in \left[\sqrt{\frac{2}{\pi}}, 1\right].\end{array}$$

*If $A$ is diagonal, or has nonnegatve entries, then the bound $\Omega_{p,r}(A)$ coincides with $\|A\|_{p,r}$.*

$$\infty \geq p \geq 2 \geq r \geq 1 \Rightarrow \quad \|A\|_{p,r} \leq \Omega_{p,r}(A) \leq \Theta(p,r,n)\|A\|_{p,r}$$

♡ **The factor $\Theta(p,r,n)$ is as follows:**

- **If either $p$ remains bounded away from $\infty$: $p \leq \widehat{p} < \infty$, or $r$ remains bounded away from $1$: $r \geq \hat{r} > 1$, or both, $\Theta(p,r,n)$ remains bounded as $n \to \infty$. E.g.,**

$$\max[\Theta(p,2,n), \Theta(2,r,n)] \leq \sqrt{\pi/2} = 1.253...$$
$$p \leq 11 \Rightarrow \Theta(p,r,n) \leq 2.6$$
$$r \geq 1.1 \Rightarrow \Theta(p,r,n) \leq 2.6$$

- **As $p \to 2+0$, $r \to 2-0$, $\Theta(p,r,n) \to 1$ uniformly in $n$. E.g.,**

$$2.25 \geq p \geq 2 \geq r \geq 1.75 \Rightarrow \Theta(p,r,n) \leq 1.1$$

- **$\Theta(p,r,n)$ admits a "nearly dimension-independent" upper bound:**

$$\Theta(p,r,n) \leq \Theta(\infty,1,n) = \sqrt{\pi \ln(n+1)}.$$

**E.g.,**

$$n \leq 10^6 \Rightarrow \Theta(p,r,n) \leq 6.6.$$

# Sketch of the Proof

♣ **To be concrete, let** $p=4$**,** $r=\frac{3}{2}$**. Then**

$$\Omega=\max_X\left\{\left(\sum_{i=1}^n[a_i^TXa_i]^{\frac{3}{4}}\right)^{\frac{2}{3}}:\ \begin{array}{l}\underbrace{\|d(X)\|_2\leq 1}_{(a)}\\ X\succeq 0\end{array}\right\}\Rightarrow X_*.$$

**Let** $\zeta\sim\mathcal{N}(0,X_*)$**. Then**

$$\|A\|_{4,\frac{3}{2}}^2\|\xi\|_4^2\geq\|A\zeta\|_{\frac{3}{2}}^2\Rightarrow$$

$$\|A\|_{4,\frac{3}{2}}^2\mathbf{E}\left\{\|\zeta\|_4^2\right\}\geq\mathbf{E}\left\{\|A\zeta\|_{\frac{3}{2}}^2\right\}=\mathbf{E}\left\{\left(\sum_i|a_i^T\zeta|^{\frac{3}{2}}\right)^{\frac{4}{3}}\right\}$$

$$\geq(\sum_i\underbrace{\mathbf{E}\left\{|a_i^T\zeta|^{\frac{3}{2}}\right\}}_{\alpha(a_i^TX_*a_i)^{\frac{3}{4}}})^{\frac{4}{3}}=\alpha^{\frac{4}{3}}\underbrace{\left(\sum_i[a_i^TX_*a_i]^{\frac{3}{4}}\right)^{\frac{4}{3}}}_{\Omega^2}$$

**and**

$$\begin{array}{l}\mathbf{E}\left\{\|\zeta\|_4^2\right\}\\ =\mathbf{E}\left\{\left(\sum_i\zeta_i^4\right)^{\frac{1}{2}}\right\}\leq\left(\mathbf{E}\left\{\sum_i\zeta_i^4\right\}\right)^{\frac{1}{2}}=\beta\underbrace{\left(\sum_i(X_*)_{ii}^2\right)^{\frac{1}{2}}}_{\leq 1\text{ by }(a)},\end{array}$$

**whence**

$$\alpha^{\frac{4}{3}}\Omega^2\leq\beta\|A\|_{4,\frac{3}{2}}^2\Rightarrow\boxed{\Omega\leq[\beta^{\frac{1}{2}}\alpha^{-\frac{2}{3}}]\|A\|_{4,\frac{3}{2}}}$$

$$\boxed{\Omega \leq [\beta^{\frac{1}{2}}\alpha^{-\frac{2}{3}}]\|A\|_{4,\frac{3}{2}}}$$

**Computing $\alpha$ and $\beta$, we arrive at**

$$\Omega_{4,\frac{3}{2}}(A) \leq \frac{\Phi(4,n)}{\Psi(3)} \qquad [p = 4, r_* = 3]$$

**Similar computation as applied to $\Omega_{3,\frac{4}{3}}(A^T) = \Omega_{4,\frac{3}{2}}(A)$ yields**

$$\Omega_{4,\frac{3}{2}}(A) \leq \frac{\Phi(3,n)}{\Psi(4)}\|A\|_{4,\frac{3}{2}},$$

**and finally**

$$\Omega_{4,\frac{3}{2}}(A) \leq \min\left[\frac{\Phi(4,n)}{\Psi(3)}, \frac{\Phi(3,n)}{\Psi(4)}\right] \cdot \|A\|_{4,\frac{3}{2}}.$$