From Estimating Linear Forms to Conic Representations of Convex-Concave Functions and Monotone Vector Fields

Arkadi Nemirovski
Georgia Institute of Technology
joint research with
Anatoli Juditsky
Université Grenoble Alpes

Optimization Without Borders
a.k.a.
Nesterov 65 & **Protasov 50**
Sochi Sirius July 14 2021

Dear Yura & Volodya,
Happy Birthdays! Till 120!

Disclaimer: *This talk is not about new algorithms of Convex Optimization. It is about extending the scope of existing algorithms.*

Full version: A. Juditsky, A. Nemirovski, *On well-structured convex-concave saddle point problems and variational inequalities with monotone operators.* `https://arxiv.org/pdf/2102.01002.pdf`, to appear in *Optimization Methods and Software*

Advertising Example

Statistical problem: *Given a polyhedral subset $\mathcal{X}=\{x: Rx\leq r\}$ of n-dimensional probabilistic simplex, $m\times n$ column-stochastic matrix $\mathcal{A}$, linear form $g(z)=g^Tz$ on $\mathbb{R}^n$, and K i.i.d. observations $\omega_k\sim\mathcal{A}x_*$ stemming from unknown signal $x_*\in\mathcal{X}$, we want to recover g^Tx_**

Fact [Statistics]: Near-optimal in the minimax sense recovery of g^Tx_* reduces to solving the convex (and thus efficiently solvable) problem

$$\min_{\alpha>0,\phi}\left\{\max_{x:Rx\leq r,y:Ry\leq r}\frac{1}{2}\left[\alpha\ln\left(\sum_i e^{\phi_i/\alpha}[\mathcal{A}x]_i\right)+\alpha\ln\left(\sum_i e^{-\phi_i/\alpha}[\mathcal{A}y]_i\right)+g^T[y-x]\right]+C\alpha\right\}$$

But: implicit nature of the objective—presence of $\max_{x:Rx\leq r,y:Ry\leq r}$—prevents utilizing highly efficient and reliable "off the shelf" convex programming solvers.

- Applying techniques to be outlined in the talk, the problem can be rewritten equivalently, *in a systematic fashion*, as

$$\min_{\alpha>0,\phi,\lambda^\pm,u_\pm,\mu_\pm,\xi^\pm,\eta^\pm}\left\{\frac{1}{2}[u_++u_-]+C\alpha:\left\{\begin{array}{l}\xi^+\geq\lambda^+,\eta^+\geq 0,\xi^-\geq\lambda^-,\eta^-\geq 0;\\ R^T\eta^+-\mathcal{A}^T\xi^+=-g,r^T\eta^+\leq\alpha-\mu_++u_+,\\ R^T\eta^--\mathcal{A}^T\xi^-=g,r^T\eta^-\leq\alpha-\mu_-+u_-;\\ \phi_i-\mu_++\alpha\ln(\alpha/\lambda_i^+)\leq 0,\ -\phi_i-\mu_-+\alpha\ln(\alpha/\lambda_i^-)\leq 0,\ \forall i\end{array}\right.\right\}.$$

The reformulated problem possesses explicitly given objective and constraints and can be fed "as is" to "off the shelf" software like `CVX`.

Well-Structured Convex Problems

♣ A convex program $\min_{x \in X} f(x)$ always has a lot of a priori known structure – otherwise, how could you know that the problem is convex?

♡ A good algorithm, in contrast to black-box-oriented "universal" algorithms like the Ellipsoid Method, should utilize a priori knowledge of the structure (think about Simplex Method fully adjusted to LP!)

♠ **However:** "structure" has no formal definition: we recognize it (say, in LP) only *after* we see it...

♣ One of the outcomes of Interior Point Revolution was discovering a specific "structure-revealing" reformulation of a convex problem – *conic formulation*

$$\min_x \left\{c^T x : Ax - b \in \mathbf{K}\right\}$$

where $\mathbf{K}$ is

—in theory, a regular (closed convex pointed with nonempty interior) cone (mo more "structure-revealing" than in the standard Mathematical Programming formulation)

—in optimization practice, a cone from Magic Family comprised of direct products of (a) nonnegative rays $\mathbb{R}_+$, (b) Lorentz cones $\mathbf{L}^n = \{x \in \mathbb{R}^n : x_n \geq \sqrt{\sum_{i=1}^{n-1} x_i^2}\}$, and (c) semidefinite cones $\mathbf{S}^n_+ = \{X = X^T \in \mathbb{R}^{n \times n} : u^T X u \geq 0\, \forall u\}$.

• Conic problems from Magic family can be solved to high accuracy in few iterations by "unified" interior point polynomial time algorithms with steadily improving "off the shelf" implementations.

♠ *For all practical purposes,* Convex Optimization is within the grasp of Magic Conic Programming. Reducing problem of interest to a magic one requires utilizing, on a case-by-case basis, a priori knowledge of problem's structure, whatever it means.

♣ **Reduction** of problem of interest, given in the Mathematical Programming "maiden form" as

$$\min_x \{f_0(x) : f_i(x) \le 0, i \le m, x \in X\} \qquad (MP)$$

to a Magic problem is immediate when we have at our disposal *semidefinite representations* of objective, constraints, and the domain of (MP).

♠ *Given family of regular cones $\mathcal{C}$ containing nonnegative rays and closed w.r.t. taking finite direct products and passing from a cone to its dual*,

— *$\mathcal{C}$-representation of a set $X \subset \mathbb{R}^n$* is

$$X = \{x : \exists u : Ax + Bu - c \in \mathbf{K}\} \text{ with } \mathbf{K} \in \mathcal{C}$$

– as *affine image* of *inverse affine image* of a cone from $\mathcal{C}$.

— *$\mathcal{C}$-representation of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$* is $\mathcal{C}$-representation

$$t \ge f(x) \;\Leftrightarrow\; \exists u : A[x;t] + Bu - c \in \mathbf{K} \text{ with } \mathbf{K} \in \mathcal{C}$$

of the epigraph $\mathrm{Epi}\{f\} = \{[x;t] : t \ge f(x)\}$ of f.

♠ **Observation:** *Given $\mathcal{C}$-representations of components of (MP):*

$$\{t \ge f_i(x) \Leftrightarrow \exists u_i : A_i[x;t] + B_i u_i - c_i \in \mathbf{K}_i\}, 0 \le i \le m \;\&\; X = \{x : \exists u : Ax + Bu - c \in \mathbf{K}\}$$

(MP) can be immediately reformulated as the conic problem

$$\min_{t,u,\{u_i\}} \left\{ t : \begin{bmatrix} A_0[x;t] + B_0 u_0 - c_0 \\ \cdots \\ A_m[x;0] + B_m u_m - c_m \\ Ax + Bu - c \end{bmatrix} \in \underbrace{\mathbf{K}_0 \times \mathbf{K}_1 \times \ldots \times \mathbf{K}_m \times \mathbf{K}}_{\in \mathcal{C}} \right\}$$

on a cone from $\mathcal{C}$.

♠ **Reduction** of convex problem of interest to $\mathcal{C}$-conic problem heavily utilizes *calculus of $\mathcal{C}$-Representations* ($\mathcal{CR}$'s) *of functions and sets.*

♡ **Raw Materials of Calculus** do depend on $\mathcal{C}$ and are comprised by "elementary" functions/sets with $\mathcal{C}$-representations obtained on case-by-case basis by bare hands, e.g.

- halfspace:

$$\{[x;t] : t \geq a^T x\} = \{[x;t] : t - a^T x \in \mathbb{R}_+\}$$

- hypograph of geometric mean, assuming $\mathcal{C}$ contains Lorentz cones:

$$\{x \geq 0, y \geq 0, t \leq \sqrt{xy}\}$$
$$\Updownarrow$$
$$\left\{[x;y;t] : \exists \tau : \underbrace{x \geq 0, y \geq 0, \tau \geq t, \sqrt{4\tau^2 + (x-y)^2} \leq x + y}_{\Leftrightarrow [[x;y;\tau-t];[2\tau;x-y;x+y]] \in \mathbb{R}^3_+ \times \mathbf{L}^3}\right\}$$

- sum of k leading eigenvalues of symmetric matrix, assuming $\mathcal{C}$ contains semidefinite cones:

$$\{t \geq \sum_{i=1}^k \lambda_i(X)\} \Leftrightarrow \{\exists Z, s : X \preceq Z + sI_n, Z \succeq 0, \mathrm{Tr}(Z) + ks \leq t\}$$
$$[\lambda_1(X) \geq \lambda_2(X) \geq \lambda_n(X) \text{ are eigenvalues of } X \in \mathbf{S}^n]$$

♡ **Calculus rules** are simple, fully algorithmic and *completely independent of what $\mathcal{C}$ is.* These rules say that *all basic convexity-preserving operations with functions and sets, as applied to $\mathcal{C}$-representable operands, produce $\mathcal{C}$-representable results with $\mathcal{C}$-representations readily given by those of the operands.* This includes

♠ For sets:

- taking finite intersections, direct products, arithmetic sums, affine and inverse affine images
- taking convex hulls of finite unions†, conic hulls†, polars†
- passing from a set to its recessive cone† and to its support function†

♠ For functions:

- taking linear combinations with nonnegative coefficients, affine substitution of variables, projective transformation
- direct summation $f_i(x_i), i = 1,..,m \mapsto f(x_1,...,x_m) = \sum_i f_i(x_i)$
- passing from functions $f_1,...,f_m,F$ to their composition $F(f_1(\cdot),...,f_m(\cdot))$, under standard monotonicity assumptions ensuring convexity of the composition
- partial minimization†, passing from function to its Fenchel conjugate †

♡ *Operations marked † require mild regularity assumptions like closedness/compactness of some operands and/or essentially strict feasibility of their $\mathcal{C}$-representations*

♠ **Note:** Fully algorithmic calculus of $\mathcal{C}$-representations can be built into a compiler. For the family of Magic cones, this is what is done by `CVX`

Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming

`http://cvxr.com/cvx`

• `CVX` is second-to-none in terms of its scope and user-friendliness "go-between" for processing well-structured convex problems reduced to (or well approximated by) Linear/Second Order Conic/Semidefinite Programming.

• `CVX` gets on input high level description of objective and constraints and uses calculus of Magic conic representations (a.k.a. SDR's - *SemiDefinite Representations*) to recognize that subsequent steps in this description are covered by calculus (this is where *disciplined* comes from). If it is the case, `CVX` automatically applies calculus rules to end up with semidefinite reformulation of the problem and sends the resulting "standard form" SDP to SDP solver.

• The solution found by the solver is then "transformed back" to the original "problem language" and returned to the user.

♠ CVX is extremely user-friendly.

Example: Inscribing largest volume ellipsoid into polytope $\{x : Ax \le b\}$.

- **Human formulation:** *Given* $m \times n$ *matrix* A *with rows* a_i^T *and* $b \in \mathbb{R}^m$, *maximize* $\mathrm{Det}(X)$ *over* $X \in \mathbf{S}^n_+$ *and* $c \in \mathbb{R}^n$ *such that* $\|Xa_i\|_2 + a_i^T c \le b_i,\ i \le i \le m$.

Explanation: We represent a candidate ellipsoid as $E = \{c + Xu : \|u\|_2 \le 1\}$ with $X \succeq 0$. The constraints on X and c state that $E \subset \{x : Ax \le b\}$, and $\mathrm{Det}(X)$ is proportional to the volume of E.

- **CVX formulation:**

```
[m,n]=size(A)
cvx_begin
variable c(n,1)
variable X(n,n) symmetric
X == semidefinite(n)
for i=1:n
  norm(A(i,:)*X)+A(i,:)*c <= b(i)
end
maximize det_rootn(X)
cvx_end
```

Note: CVX is enough intelligent to know SDR of $-\mathrm{Det}^{1/n}(X)$, $X \in \mathbf{S}^n_+$ (`-det_rootn(X)` in CVX), same as SDR's of tens of other useful functions.

♣ **Our goal:** *To define conic representations of "well-structured" convex-concave functions and monotone vector fields and to develop calculus of these representations, with the ultimate goal to reduce the associated Saddle Point problems and Variational Inequalities to conic programs.*

♠ **Convention:** *From now on we fix a family $\mathcal{C}$ of regular cones in Euclidean spaces which contains nonnegative rays, $\mathbf{L}^3$, and is closed w.r.t. taking finite direct products and passing from a cone $\mathbf{K}$ to its dual $\mathbf{K}^*$.*

(!) Unless otherwise is explicitly stated, all cones below belong to $\mathcal{C}$.

Terminology: A conic constraint $Ax - b \in \mathbf{K}$ is called *essentially strictly feasible (e.s.f.)*, if the cone $\mathbf{K}$ can be represented as $\mathbf{K} = \mathbf{M} \times \mathbf{P}$ with regular $\mathbf{M}$ and polyhedral $\mathbf{P}$ in such a way that

$$\exists \overline{x} : A\overline{x} - b \in [\mathrm{int}\, \mathbf{M}] \times \mathbf{P}.$$

Conic Representations of Convex-Concave Functions [Jud&N.'21]

Definition. *Let $\mathcal{X}$, $\mathcal{Y}$ be nonempty convex sets given by $\mathcal{C}$-representations:*

$$\mathcal{X} = \{x : \exists \xi : Ax + B\xi - c \in \mathbf{K}_{\mathcal{X}}\},\ \mathcal{Y} = \{y : \exists \eta : Cy + D\eta - e \in \mathbf{K}_{\mathcal{Y}}\}$$

and let $\psi(x;y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a continuous convex-concave function. ψ is called $\mathcal{C}$-representable ($\mathcal{C}r$) on $\mathcal{X} \times \mathcal{Y}$, if it admits a $\mathcal{C}$-representation ($\mathcal{CR}$):

$$\forall (x \in \mathcal{X}, y \in \mathcal{Y}) : \psi(x;y) = \inf_{f,t,u} \left\{ f^T y + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$

We call the above $\mathcal{CR}$ essentially strictly feasible (e.s.f.), if the conic constraint

$$Pf + tp + Qu + [Rx - s] \in \mathbf{K}$$

in variables f, t, u is e.s.f. for every $x \in \mathcal{X}$.

Motivation: In the situation of Definition, the set

$$\Psi = \{[x; f; t] : x \in \mathcal{X}, f^T y + t \geq \psi(x;y)\, \forall y \in \mathcal{Y}\} \qquad (!)$$

is convex, and by usual Fenchel duality in y-variable, one has

$$\psi(x;y) = \inf_{f,t} \left\{ f^T y + t : (x, f, t) \in \Psi \right\}\ \forall (x,y) \in \mathcal{X} \times \mathcal{Y} \qquad (!!)$$

$\mathcal{CR}$ is obtained from (!!) by replacing Ψ (which by itself can be too complicated to be $\mathcal{C}r$) with a (perhaps, smaller) $\mathcal{C}r$ set which still ensures (!!).

Note: The convex set Ψ is a natural candidate to the role of the *epigraph* of convex-concave function ψ on $\mathcal{X} \times \mathcal{Y}$ – look what happens when $\mathcal{Y} = \{0\}$.

$$\mathcal{X} = \{x : \exists \xi : Ax + B\xi - c \in \mathbf{K}_{\mathcal{X}}\},\ \mathcal{Y} = \{y : \exists \eta : Cy + D\eta - e \in \mathbf{K}_{\mathcal{Y}}\}$$
$$\forall (x \in \mathcal{X}, y \in \mathcal{Y}) : \psi(x; y) = \inf_{f,t,u} \left\{ f^T y + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$

♠ **Main Observation:** *Assume $\mathcal{Y}$ is compact with e.s.f. $\mathcal{CR}$. Then the problem*

$$\min_{x \in \mathcal{X}} \left\{ \overline{\psi}(x) = \max_{y \in \mathcal{Y}} \psi(x; y) \right\} \qquad (P)$$

reduces to the explicit $\mathcal{C}$-conic problem

$$\min_{x,\xi,f,t,u,\lambda} \left\{ t - e^T\lambda : \begin{array}{c} Pf + tp + Qu + Rx - s \in \mathbf{K} \\ C^T\lambda + f = 0, D^T\lambda = 0, \lambda \in \mathbf{K}_{\mathcal{Y}}^* \\ Ax + B\xi - c \in \mathbf{K}_{\mathcal{X}} \end{array} \right\} \qquad (Q)$$

"reduces" meaning that the x-component of a feasible solution $\zeta = (x, \xi, f, t, u, \lambda)$ to (Q) is a feasible solution to (P) with the value of the objective of (P) at x being $\leq$ the value of the objective of (Q) at ζ, and the optimal values in (P) and (Q) are the same.

$\Rightarrow$ *For every $\epsilon > 0$, every feasible ϵ-optimal approximate solution to (Q) induces feasible ϵ-optimal approximate solution to (P).*

$$\mathcal{X} = \{x : \exists \xi : Ax + B\xi - c \in \mathbf{K}_{\mathcal{X}}\},\ \mathcal{Y} = \{y : \exists \eta : Cy + D\eta - e \in \mathbf{K}_{\mathcal{Y}}\}$$
$$\forall (x \in \mathcal{X}, y \in \mathcal{Y}) : \psi(x; y) = \inf_{f,t,u} \left\{ f^T y + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$

Reason for Main Observation:

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x; y) = \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \inf_{f,t,u} \left\{ f^T y + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$
$$= \min_{x \in \mathcal{X}} \inf_{f,t,u} \left\{ \left[\max_{y \in \mathcal{Y}} f^T y \right] + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$
$$= \inf_{x \in \mathcal{X}, f, t, u} \left\{ \left[\max_{y,\eta} \left\{ f^T y : Cy + D\eta - e \in \mathbf{K}_{\mathcal{Y}} \right\} \right] + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$
$$= \inf_{x \in \mathcal{X}, f, t, u} \left\{ \left[\min_{\lambda \in \mathbf{K}_{\mathcal{Y}}^*} \left\{ -e^T \lambda : C^T \lambda + f = 0, D^T \lambda = 0 \right\} \right] + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$

[by Conic Duality]

$$= \inf_{x \in \mathcal{X}, \lambda \in \mathbf{K}_{\mathcal{Y}}^*, f, t, u} \left\{ t - e^T \lambda : \begin{array}{l} Pf + tp + Qu + Rx - s \in \mathbf{K} \\ C^T \lambda + f = 0, D^T \lambda = 0 \end{array} \right\}$$

♠ **Fact** ["Symmetry"]: *Essentially strictly feasible* $\mathcal{CR}$

$$\forall (x \in \mathcal{X}, y \in \mathcal{Y}) : \psi(x; y) = \inf_{f,t,u} \left\{ f^T y + t : Pf + tp + Qu + Rx - s \in \mathbf{K} \right\}$$

of continuous convex-concave function $\psi(x; y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ *induces explicit* $\mathcal{CR}$

$$\forall (\widetilde{x} \in \widetilde{\mathcal{X}}) : \widetilde{\psi}(\widetilde{x}, \widetilde{y}) = \inf_{\widetilde{f}, \widetilde{t}, \widetilde{u}} \left\{ \widetilde{f}^T \widetilde{y} + \widetilde{t} : \begin{array}{l} \widetilde{f} = R^T \widetilde{u}, \widetilde{t} + s^T \widetilde{u} = 0, Q^T \widetilde{u} = 0, \\ p^T \widetilde{u} = 1, P^T \widetilde{u} = \widetilde{x}, \widetilde{u} \in \mathbf{K}^* \end{array} \right\}$$

of the "symmetric entity" – convex-concave function

$$\widetilde{\psi}(\widetilde{x}, \widetilde{y}) := -\psi(\widetilde{y}; \widetilde{x}) : \widetilde{\mathcal{X}} \times \widetilde{\mathcal{Y}} := \mathcal{Y} \times \mathcal{X} \to \mathbb{R}.$$

implying, in the case of compact $\mathcal{X}$ *with e.s.f.* $\mathcal{CR}$, *explicit* $\mathcal{C}$*-conic reformulation of the problem*

$$\max_{y \in \mathcal{Y}} \left[\underline{\psi}(y) := \min_{x \in \mathcal{X}} \psi(x; y) \right]$$

Calculus of $\mathcal{C}r$ Functions: "Generic" Raw materials

♣ *The following continuous convex-concave functions $\psi(\cdot;\cdot):\mathcal{X}\times\mathcal{Y}\to\mathbb{R}$ are $\mathcal{C}r$ with explicit $\mathcal{CR}$'s:*

- $\psi(x;y)=a(x)$ where $a(x)$, $\mathrm{Dom}\,a\supset\mathcal{X}$, is given by explicit $\mathcal{CR}$
- $\psi(x;y)=-b(y)$ where $b(y)$, $\mathrm{Dom}\,b\supset\mathcal{Y}$, is given by explicit e.s.f. $\mathcal{CR}$
- Bilinear functions $\psi(x;y)\equiv a^Tx+b^Ty+x^TAy+c$
- **Generalized bilinear functions.** Let $\mathbf{U}\in\mathcal{C}$, and E be embedding Euclidean space of the cone $\mathbf{U}$. Then the following functions are $\mathcal{C}r$ with explicit $\mathcal{CR}$'s:

a) functions of the form $\psi(x;y)=\langle F(x),y\rangle:\mathcal{X}\times\mathbf{U}^*\to\mathbb{R}$, where
- $\mathcal{X}$ is given by explicit $\mathcal{C}$R
- $F(x):\mathcal{X}\to E$ is continuous with $\mathbf{U}$-epigraph $\{(x,U):U-F(x)\in\mathbf{U}\}$ given by explicit $\mathcal{C}$R.

b) functions of the form $\psi(x;u)=\langle x,G(y)\rangle:\mathbf{U}\times\mathcal{Y}\to\mathbb{R}$, where
- $\mathcal{Y}$ is given by explicit $\mathcal{C}$R
- $G(y):\mathcal{Y}\to E$ is continuous with $\mathbf{U}^*$-hypograph $\{(y,U):G(y)-U\in\mathbf{U}^*\}$ given by explicit e.s.f. $\mathcal{C}$R

c) functions of the form $\psi(x;y)=\langle F(x),G(y)\rangle:\mathcal{X}\times\mathcal{Y}\to\mathbb{R}$, where
- $\mathcal{X}$ and $\mathcal{Y}$ are given by explicit $\mathcal{C}$R's
- $F(x):\mathcal{X}\to\mathbf{U}$ is continuous with $\mathbf{U}$-epigraph given by explicit $\mathcal{C}$R
- $G(y):\mathcal{Y}\to\mathbf{U}^*$ is continuous with $\mathbf{U}^*$-hypograph given by explicit e.s.f. $\mathcal{C}$R

Illustration: Let $\mathcal{C}$ contain semidefinite cones, $\mathcal{X} = \mathbb{R}^{m\times n}$, $\mathcal{Y} = \mathbf{S}^n_+$. The function

$$\psi(x;y) = \mathrm{Tr}(x^T x y^{1/2}) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$$

is a Generalized Bilinear function of type c) given by the data

$$\mathbf{U} = \mathbf{S}^n_+,\ F(x) = x^T x,\ G(y) = y^{1/2}.$$

The corresponding Calculus rule results in

$$\begin{array}{l}\forall (x \in \mathbb{R}^{m\times n}, y \in \mathbf{S}^n_+):\\ \psi(x;y) = \inf\limits_{f,t,u=(z,\beta,\gamma)} \left\{ \mathrm{Tr}(fy) + t : \left\{ \begin{array}{l} f \in \mathbf{S}^n, \beta \in \mathbb{R}^{n\times n}, \gamma \in \mathbf{S}^n, z \in \mathbf{S}^n \\ t = \mathrm{Tr}(\gamma), z \preceq \beta + \beta^T \\ \left[\begin{array}{c|c} f & \beta \\ \hline \beta^T & \gamma \end{array}\right] \succeq 0, \left[\begin{array}{c|c} z & x^T \\ \hline x & I_m \end{array}\right] \succeq 0 \end{array} \right\} \right\}.\end{array}$$

Calculus of $\mathcal{C}r$ Functions: "Ad Hoc" Raw materials

A. The function

$$\psi(x;y) = -\frac{x}{x+y+1} : \underbrace{\mathbb{R}_+}_{\mathcal{X}} \times \underbrace{\mathbb{R}_+}_{\mathcal{Y}} \to \mathbb{R}$$

admits $\mathcal{CR}$

$$\psi(x;y) = \min_{f,t,u} \left\{ fy + t : \quad u \geq 0, \left[\begin{array}{c|c} x & s \\ \hline s & f \end{array}\right] \succeq 0, \left[\begin{array}{c|c} t-f+1 & 1-s \\ \hline 1-s & 1 \end{array}\right] \succeq 0 \right\}.$$

B. Let $\mathcal{X} \subset \mathbb{R}^m$, $\mathcal{Y} \subset \mathbb{R}^m_+ \backslash \{0\}$ are $\mathcal{C}r$, and $\mathcal{Y}$ be compact. The function

$$\psi(x; y) = \ln\left(\sum_{i=1}^m \mathrm{e}^{x_i} y_i\right) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$$

admits representation

$$\psi(x; y) = \inf_{f,t,u} \left\{ f^T y + t : f_i \geq \mathrm{e}^{x_i + u} \,\forall i \;\&\; t \geq -u - 1 \right\}$$

This is $\mathcal{CR}$, provided that $\mathcal{C}$ contains the exponential cone

$$\mathbf{E} = \mathrm{cl}\{[t; s; r] : t \geq s\mathrm{e}^{r/s} \;\&\; s > 0\}.$$

Note: "For all practical purposes," $\mathbf{E}$ is SDr. Formally: *conic constraints involving* $\mathbf{E}$ *are polynomially reducible to semidefinite (in fact, even linear) constraints.*

♠ Here is the $\mathcal{CR}$ of a function $E(x)$ approximating $\exp\{x\}$ in the range $-700 \leq x \leq 700$ (the range of exponent which "lives in computer") within relative error $\leq$ 3.e-11:

$$t \geq E(x) := \left[\sum_{\ell=0}^{6} \frac{(x/2^{15})^\ell}{\ell!}\right]^{[2^{15}]}$$

$$\Updownarrow$$

$$\exists u_0, u_1, u_2, u_3, v, \tau_1, \tau_2, \tau_3, s, w_1, ..., w_{14} :$$

$$-700 \leq x \leq 700$$

$$\frac{x}{32768} + 1 \leq u_0,\ 0 \leq u_1 \leq \sqrt{\tau_1 u_0},\ 0 \leq u_2 \leq \sqrt{u_0},\ 0 \leq u_3 \leq \sqrt{u_1 u_2},\ u_0 \leq \sqrt{u_3}$$

$$\left(\frac{x}{32768} + \frac{5}{3}\right)^2 \leq v,\ v^2 \leq \tau_2,\ \left(\frac{x}{32768} + \frac{1963}{855}\right)^2 \leq \tau_3,\ s \geq \frac{78871}{5540400} + \frac{19\tau_3}{144} + \frac{\tau_2}{48} + \frac{\tau_1}{720}$$

$$w_1 \geq s^2,\ w_2 \geq w_1^2,\ ...,\ w_{14} \geq w_{13}^2,\ t \geq w_{14}^2$$

C. Let $p>1$, $\mathcal{X}\subset\mathbb{R}^m$ and $\mathcal{Y}\subset\mathbb{R}^m_+$ be $\mathcal{C}r$, and $\theta_i(x):\mathcal{X}\to\mathbb{R}_+$, $1\leq i\leq m$, be $\mathcal{C}r$. Then the function

$$\psi(x;y)=\left(\sum_{i=1}^m \theta_i^p(x)y_i\right)^{1/p}:\mathcal{X}\times\mathcal{Y}\to\mathbb{R}$$

admits representation

$$\psi(x;y)=\inf_{[f;t]}\left\{f^Ty+t:t\geq 0, f\geq 0, t^{\frac{p-1}{p}}f_i^{\frac{1}{p}}\geq\kappa\theta_i(x), i\leq m\right\}$$
$$\left[\kappa=p^{-1}(p-1)^{\frac{p-1}{p}}\right]$$

which is $\mathcal{CR}$, provided p is rational.

D. Let $\mathcal{X} \subset \mathbb{R}^{m\times n}$ and $\mathcal{Y} \subset \mathbf{S}^m_+$ be SDr. Then the function

$$\psi(x;y) = 2\sqrt{\mathrm{Tr}(x^T y x)} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$$

admits SDr

$$\forall (x \in \mathcal{X}, y \in \mathcal{Y}) : \psi(x;y) = \inf_{f,t} \left\{ \mathrm{Tr}(yf) + t : \left[\begin{array}{c|c} f & x \\ \hline x^T & tI_n \end{array}\right] \succeq 0 \right\}$$

This is how **D** works in Robust Markowitz Portfolio Selection

$$\min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \left[-r^T x + 2\mu\sqrt{x^T y x} \right] \qquad (M)$$

where $x \in \mathbb{R}^n$ is composition of portfolio, r is the vector of expected returns, $\mu > 0$ is "safety parameter," y is the uncertain covariance matrix of the returns running through the set

$$\mathcal{Y} = \{ y \in \mathbf{S}^n_+ : \sum_\tau [a_{i\tau}^T y b_{i\tau} + b_{i\tau} y a_{i\tau}^T] \preceq p_i, i \le I, y_- \le y \le y_+ \} \qquad [\le \text{ acts entrywise}]$$

(M) is equivalent to the explicit SDP

$$\min_{x,s,\{\alpha_i\},M_\pm} \left\{ \begin{array}{l} -r^T x + \mu \left[s + \sum_i \mathrm{Tr}(\alpha_i p_i) + \mathrm{Tr}(M_+ y_+ - M_- y_-) \right] : \\ \left\{ \begin{array}{l} \left[\begin{array}{c|c} \sum_i \sum_\tau \left[a_{i\tau} \alpha_i b_{i\tau}^T + b_{i\tau} \alpha_i a_{i\tau}^T \right] + M_+ - M_- & x \\ \hline x^T & s \end{array}\right] \succeq 0 \\ \alpha_i \succeq 0, i \le I, M_\pm \ge 0, x \in \mathcal{X} \end{array} \right. \end{array} \right\}$$

Calculus of $\mathcal{C}r$ Functions: Calculus rules

♣ *The following operations with $\mathcal{C}r$ operands yield $\mathcal{C}r$ convex-concave functions with $\mathcal{CR}$'s readily given by $\mathcal{CR}$'s of the operands:*

- Direct summation:

$$\{\psi_i(x_i;y_i):\mathcal{X}_i\times\mathcal{Y}_i\to\mathbb{R},\theta_i>0\}_{i\leq K}$$
$$\Rightarrow \psi(x_1,...,x_K;y_1,...,y_K)=\sum_i\theta_i\psi_i(x_i;y_i):[\mathcal{X}_1\times...\times\mathcal{X}_K]\times[\mathcal{Y}_1\times...\times\mathcal{Y}_K]\to\mathbb{R}$$

- Taking conic combinations:

$$\{\psi_i(x;y):\mathcal{X}\times\mathcal{Y}\to\mathbb{R},\theta_i>0\}_{i\leq K}$$
$$\Rightarrow \psi(x;y)=\sum_i\theta_i\psi_i(x;y):\mathcal{X}\times\mathcal{Y}\to\mathbb{R}$$

- Affine substitution of variables:

$$\{A(x)=Ax+a:\mathcal{X}\to\mathcal{X}^+,B(y)=By+b:\mathcal{Y}\to\mathcal{Y}^+,\psi_+(\xi;\eta):\mathcal{X}^+\times\mathcal{Y}^+\to\mathbb{R}\}$$
$$\Rightarrow \psi(x;y)=\psi_+(A(x);B(y)):\mathcal{X}\times\mathcal{Y}\to\mathbb{R}$$

- Projective transformation in x-variable:

$$\{\psi(x;y):\mathcal{X}\times\mathcal{Y}\to\mathbb{R}\}$$
$$\Rightarrow \mathcal{X}_+:=\{(x,\alpha):\alpha>0,\alpha^{-1}x\in\mathcal{X}\}\ \&\ \psi_+((x,\alpha);y):=\alpha\psi(x/\alpha;y):\mathcal{X}^+\times\mathcal{Y}\to\mathbb{R}$$

♣ *The following operations with $\mathcal{C}r$ operands yield $\mathcal{C}r$ convex-concave functions with $\mathcal{CR}$'s readily given by $\mathcal{CR}$'s of the operands:*

- Superposition in x-variable:

$$\begin{array}{c} \mathcal{X}, \mathcal{Y}, \mathbb{R}^n \supset \mathcal{Z} : \mathcal{C}r, \ \mathbb{R}^n \supset \mathbf{U} \in \mathcal{C} \\ \psi_+(z;y) : \mathcal{Z} \times \mathcal{Y} : \ \mathcal{C}r \text{ and } \mathbf{U}\text{-monotone in } z: y \in \mathcal{Y}, z, z' \in \mathcal{Z}, z' - z \in \mathbf{U} \Rightarrow \psi_+(z;y) \le \psi_+(z';y) \\ X(x) : \mathcal{X} \to \mathcal{Z} : \text{with } \mathcal{C}r \ \mathbf{U}\text{-epigraph } \{[x;z] \in \mathcal{X} \times \mathcal{Z} : z - X(x) \in \mathbf{U}\} \end{array}$$

$$\Rightarrow \psi(x;y) = \psi_+(X(x);y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$$

- Partial maximization in y-variable:

$$\begin{array}{c} \mathcal{X}, \mathcal{Y}, \phi(x;y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} : \ \mathcal{C}r \\ \mathcal{Y} \in \mathbb{R}^p_w \times \mathbb{R}^q_z : \text{ compact with } \mathcal{C}r \ \mathcal{Y} = \{[w;z] : \exists v : Aw + Bz + Cv - d \in \mathbf{K}\} \end{array}$$

such that for all $[w;h] \in \mathcal{Y}$ conic constraint $Bz + Cv + [Aw - d] \in \mathbf{K}$ in variables z, v is e.s.f.

$$\max_{z:[w;z]\in\mathcal{Y}} \psi(x;[w;z]) \text{ is continuous on } \mathcal{X} \times \{\mathcal{W} := \{w : \exists z : [w;z] \in \mathcal{Y}\}\}$$

$$\Rightarrow \overline{\psi}(x;w) := \max_{z:[w;z]\in\mathcal{Y}} \psi(x;[w;z]) : \mathcal{X} \times \mathcal{W} \to \mathbb{R}$$

Note: By Symmetry, Calculus extends to projective transformation and superposition in y-variable and partial minimization in x-variable.

♣ *The following operation with $\mathcal{C}r$ operands yields $\mathcal{C}r$ convex-concave function with $\mathcal{CR}$ readily given by $\mathcal{CR}$'s of the operands:*

• Taking Fenchel conjugate:

$\mathcal{X}, \mathcal{Y}, \phi(x;y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$: $\mathcal{C}r$ given by e.s.f. $\mathcal{CR}$'s, with compact $\mathcal{X}, \mathcal{Y}$ and with e.s.f. conic constraint $D^T\lambda = 0$ & $\lambda \geq_{\mathbf{K}_{\mathcal{Y}}^*} 0$ in variable λ

$$\Rightarrow \phi_*(p,q) = \max_{x\in\mathcal{X}} \min_{y\in\mathcal{Y}} \left[p^T x + q^T y - \phi(x;y)\right] : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$$

Illustration: We have build $\mathcal{CR}$ for

$$\psi(x;y):-\frac{x}{x+y+1}:\mathbb{R}^2_+\to\mathbb{R}.$$

By Calculus, this $\mathcal{CR}$ induces $\mathcal{CR}$'s of functions of the form

$$-\frac{\alpha(x)}{\alpha(x)+\beta(y)+1}:\mathrm{Dom}\,\alpha\times\mathrm{Dom}\,\beta\to\mathbb{R}$$

[α,β : nonnegative concave with $\mathcal{C}r$ hypographs]

Conic Representations of Monotone Vector Fields

♠ **Preliminaries.** Let $\mathcal{X} \subset E = \mathbb{R}^n$ be a nonempty convex set and $F(x) : \mathcal{X} \to E$ be a vector field. F is called *monotone* on $\mathcal{X}$, if $\langle F(x) - F(y), x - y\rangle \geq 0 \ \forall x, y \in \mathcal{X}$.
Variational Inequality induced by $F, \mathcal{X}$ reads

$$\text{Find } x_* \in \mathcal{X} \text{ such that } \langle F(x), x - x_*\rangle \geq 0 \ \forall x \in \mathcal{X} \qquad \text{VI}(F, \mathcal{X})$$

Convention: *From now on, if otherwise is not explicitly stated, vector fields $F : \mathcal{X} \to E$ in question are assumed to be monotone and continuous on their convex domains $\mathcal{X}$.* In this case,

Solutions to $\text{VI}(F, \mathcal{X})$ are exactly the points $x_* \in \mathcal{X}$ such that $\langle F(x_*), x - x_*\rangle \geq 0 \ \forall x \in \mathcal{X}$.
Solutions do exist when $\mathcal{X}$ is compact

♡ **Inaccuracy** of a candidate solution $x \in \mathcal{X}$ to $\text{VI}(F, \mathcal{X})$ is quantified by *dual gap*

$$\epsilon_{\text{VI}}[x|\mathcal{X}] = \max_{y \in \mathcal{X}} \langle F(y), x - y\rangle$$

Basic examples of monotone vector fields:
- Gradient field $F(x) = \nabla f(x)$ of C^1 convex function $f : \mathcal{X} \to \mathbb{R}$
- Vector field $F(u; v) = [\nabla_u \psi(u; v); -\nabla_v \psi(u; v)]$ stemming from C^1 convex-concave function $\psi(u; v) : \mathcal{X} := \mathcal{U} \times \mathcal{V} \to \mathbb{R}$.

In the first case, solutions to $\text{VI}(F, \mathcal{X})$ are the minimizers of f on $\mathcal{X}$, in the second – saddle point of ψ on $\mathcal{U} \times \mathcal{V}$.

Observation [Nesterov&N'94]: *Given a continuous monotone vector field $F:\mathcal{X}\to\mathcal{R}$, $\mathcal{X}\subset E=\mathbb{R}^n$, let*

$$\mathfrak{T}[F,\mathcal{X}]=\{[t;g;x]\in\mathbb{R}\times E\times E: x\in\mathcal{X}, t-\langle g,y\rangle\geq\langle F(y),x-y\rangle\ \forall y\in\mathcal{X}\}.$$

Then

A. $\mathfrak{T}[F,\mathcal{X}]$ *is a convex set which contains all triples* $[\langle F(x),x\rangle;F(x);x]$, $x\in\mathcal{X}$*; this set is closed provided* $\mathcal{X}$ *is so. Besides this,* $\mathfrak{T}[F,\mathcal{X}]$ *is* t*-monotone:*

$$[t;g;x]\in\mathfrak{T}[F,\mathcal{X}] \text{ and } t'\geq t\Rightarrow[t';g;x]\in\mathfrak{T}[F,\mathcal{X}].$$

B. *For* $\epsilon\geq 0$*, let*

$$\mathcal{X}_*(\epsilon)=\{[g;t]\in E\times\mathbb{R}:\sup_{y\in\mathcal{X}}[t-\langle g,y\rangle]\leq\epsilon\},$$

so that $\mathcal{X}_*(\epsilon)$ *is a nonempty closed convex set. Then for every* $\epsilon\geq 0$*, feasible* ϵ*-solutions to* VI$(F,\mathcal{X})$ *– points* $x\in\mathcal{X}$ *with* $\epsilon_{\rm VI}[x|\mathcal{X}]\leq\epsilon$ *– are exactly the points*

$$x:\exists(t,g):[t;g;x]\in\mathfrak{T}[F,\mathcal{X}]\ \ \&\ \ [t;g]\in\mathcal{X}_*(\epsilon). \qquad (!)$$

Note: *Were we given* $\mathcal{CR}$*'s of* $\mathfrak{T}[F,\mathcal{X}]$ *and of* $\mathcal{X}_*(\epsilon)$*,* (!) would reduce finding ϵ-solution to VI$(F,\mathcal{X})$ to solving $\mathcal{C}$-conic problem. $\mathcal{CR}$ of $\mathcal{X}_*(\epsilon)$ "is cheap" – it is readily given by e.s.f. $\mathcal{CR}$ of $\mathcal{X}$. In contrast, $\mathfrak{T}[F,\mathcal{X}]$ can be too complicated to admit explicit $\mathcal{CR}$.

In the sequel, we intend to replace in (!) the set $\mathfrak{T}[F,\mathcal{X}]$ with a smaller $\mathcal{C}r$ set which still is "rich enough" for our ultimate goal.

$F:\mathcal{X}\to E=\mathbb{R}^n$: continuous monotone vector field on convex domain

$$\Rightarrow \boxed{\begin{array}{c}\mathfrak{T}[F,\mathcal{X}]=\{[t;g;x]\in\mathbb{R}\times E\times E: x\in\mathcal{X}, t-\langle g,y\rangle\geq\langle F(y),x-y\rangle\ \forall y\in\mathcal{X}\}\\ \hline \mathcal{X}_*(\epsilon)=\{[g;t]\in E\times\mathbb{R}:\sup_{y\in\mathcal{X}}[t-\langle g,y\rangle]\leq\epsilon\}\end{array}}$$

Definition. *$\mathcal{C}$-representation of $(F,\mathcal{X})$ is a conic constraint*

$$Xx+Gg+tT+Uu-a\in\mathbf{K}$$

in variables $t\in\mathbb{R}$, $g,x\in E$, $u\in\mathbb{R}^k$ with t-monotone projection

$$\mathcal{T}:=\{[t;g;x]:\exists u: Xx+Gg+tT+Uu-a\in\mathbf{K}\}$$

of the solution set on the space of t,g,x variables satisfying the inclusions

$$\{[t;F(x);x]:t>\langle F(x),x\rangle,x\in\mathcal{X}\}\subset\mathcal{T}\subset\mathfrak{T}[F,\mathcal{X}]$$

Conic Form of Variational Inequality with $\mathcal{C}$-Representable Monotone Vector Field

Theorem [Jud&N'21] *Assume that $\mathcal{X} \subset E = \mathbb{R}^n$ is nonempty convex compact set given by e.s.f. $\mathcal{CR}$:*

$$\mathcal{X} = \{x : \exists v : Ax + Bv - b \in \mathbf{L}\}$$

and that continuous monotone vector field $F : \mathcal{X} \to E$ admits $\mathcal{CR}$ given by conic constraint

$$Xx + Gg + tT + Uu - a \in \mathbf{K}$$

in variables t, g, x, u.

Then for every $\epsilon > 0$ finding ϵ-solution to $\mathrm{VI}(F, \mathcal{X})$ reduces to finding feasible solution to an explicit feasible conic constraint

$$Xx + Gg + tT + Uu - a \in \mathbf{K},\ A^T\lambda = g,\ B^T\lambda = 0,\ t - b^T\lambda \le \epsilon,\ \lambda \in \mathbf{L}^* \qquad (!)$$

in variables t, g, x, λ. Specifically, x-component of feasible solution to $(!)$ is a feasible ϵ-solution to $\mathrm{VI}(F, \mathcal{X})$.

Calculus of $\mathcal{CR}$'s of Monotone Vector Fields: Raw materials

♣ *The following monotone vector fields $F:\mathcal{X}\to E$ are $\mathcal{C}r$ with explicit $\mathcal{CR}$'s:*

• Affine monotone field $F(x)=Ax+b:E\to E$ $[A+A^T\succeq 0]$

• Gradient vector field $F(x)=\nabla f(x)$ of C^1 convex function $f:\mathcal{X}\to\mathbb{R}$, provided that $\mathcal{X}$ is convex compact set and the epigraph $\{(x,t):x\in\mathcal{X},t\geq f(x)\}$ of f on $\mathcal{X}$ is given by e.s.f. $\mathcal{C}r$

• Monotone vector field $F(u;v)=[\nabla_u\psi(u;v);-\nabla_v\psi(u;v)]:\mathcal{U}\times\mathcal{V}\to\mathbb{R}^{n_u+n_v}$ associated with C^1 convex-concave function ψ, provided that $\mathcal{U}$, $\mathcal{V}$ are convex compact, and $\mathcal{U},\mathcal{V},\psi$ are given by e.s.f. $\mathcal{CR}$'s

• Rational monotone *univariate* vector field $F(x)=\frac{p(x)}{q(x)}:\mathcal{X}=[0,1]\to\mathbb{R}$ (p,q are algebraic polynomials, $q>0$ on $\mathcal{X}$)

Calculus of $\mathcal{CR}$'s of Monotone Vector Fields: Calculus Rules

♣ *The results of the following monotonicity-preserving operations with $\mathcal{C}r$ monotone vector fields are $\mathcal{C}r$, with $\mathcal{CR}$'s readily given by $\mathcal{CR}$'s of operands:*

- Restriction on $\mathcal{C}$-representable set
- Direct summation
- Taking linear combinations with nonnegative coefficients
- Affine substitution of variables: $\mathcal{CR}$'s of $\mathcal{X}$ and of continuous monotone vector field $F:\mathcal{X}\to\mathbb{R}^m$ explicitly induce a $\mathcal{CR}$ of the field

$$G(y)=A^TF(Ay+b):\mathcal{Y}:=\{y\in\mathbb{R}^n:Ay+b\in\mathcal{X}\}\to\mathbb{R}^n \qquad [A:m\times n,\mathcal{Y}\neq\emptyset]$$

♠ **Illustration:** Consider *Convex Nash Equilibrium* problem, where $n \geq 2$ players make selections $x_i \in \mathcal{X}_i$, with $\mathcal{C}r$ convex compact sets $\mathcal{X}_i$. As a result of these selections, loss of i-thS player is a known C[1] function $\psi_i(x_1, ..., x_n)$, $i \leq n$.

• Nash Equilibrium $x^* \in \mathcal{X} = \mathcal{X}_1 \times ... \times \mathcal{X}_n$ is given by

$$x_i^* \in \operatorname*{Argmin}_{x_i \in \mathcal{X}_i} \psi_i(x_1^*, ..., x_{i-1}^*, x_i, x_{i+1}^*, ..., x_n^*),\ 1 \leq i \leq n$$

— a selection where no player can reduce loss by unilateral action.

• Nash Equilibrium Problem is called convex, if

a) ψ_i is convex in x_i and jointly concave in all x_j with $j \neq i$, $1 \leq i \leq n$, and

b) $\phi(x) := \sum_i \psi_i(x)$ is convex.

In this case, the vector field

$$F(x_1, ..., x_n) = [\nabla_{x_1}\psi_1(x); ...; \nabla_{x_n}\psi_n(x)]$$

is monotone, and Nash Equilibria are exactly the solutions to VI$(F, \mathcal{X})$.

♡ **Fact** [Jud&N'21] *$\mathcal{CR}$ of F is readily given by $\mathcal{CR}$'s of the monotone vector fields associated with convex-concave functions $\psi_i(\cdot)$ and with the convex function ϕ* and as such can be obtained by utilizing the related components of our Calculus.