DC Semidefinite Programming

M.V. Dolgopolik

Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences, Saint Petersburg, Russia

Variational Analysis and Optimisation Webinar
September 22, 2021

Contents

Contents

DC optimisation

DC optimisation problems:

$$\min \; f(x) = g(x) - h(x),$$

where g and h are convex functions.

DC optimisation

DC optimisation problems:

$$\min \; f(x) = g(x) - h(x),$$

where g and h are convex functions.

For any $v \in \partial h(x_0)$ one has

$$f(x) \le g(x) - h(x_0) - \langle v, x - x_0 \rangle.$$

DC optimisation

DC optimisation problems:

$$\min\ f(x) = g(x) - h(x),$$

where g and h are convex functions.

For any $v \in \partial h(x_0)$ one has

$$f(x) \le g(x) - h(x_0) - \langle v, x - x_0 \rangle.$$

If x_0 is a local minimiser of f, then it is a globally optimal solution of the convex problem

$$\min\ g(x) - \langle v, x - x_0 \rangle.$$

DC Algorithm (DCA)

Algorithm 1: DC Algorithm/The Convex-Concave Procedure (CCP).

Step 1. Choose an initial point x_0 and set $n := 0$.
Step 2. Compute $v_n \in \partial h(x_n)$.
Step 3. Set the value of x_{n+1} to a solution of the convex problem

$$\min \; g(x) - \langle v_n, x - x_n \rangle.$$

If $x_{n+1} = x_n$, **Stop**. Otherwise, put $n := n + 1$ and go to **Step 2**.

DC Algorithm (DCA)

Algorithm 2: DC Algorithm/The Convex-Concave Procedure (CCP).

Step 1. Choose an initial point x_0 and set $n := 0$.
Step 2. Compute $v_n \in \partial h(x_n)$.
Step 3. Set the value of x_{n+1} to a solution of the convex problem

$$\min \quad g(x) - \langle v_n, x - x_n \rangle.$$

If $x_{n+1} = x_n$, **Stop**. Otherwise, put $n := n + 1$ and go to **Step 2**.

Since $f(x_{n+1}) \leq f(x_n)$, the actual stopping criteria:

$$|f(x_{n+1}) - f(x_n)| < \varepsilon \quad \text{and} \quad \|x_{n+1} - x_n\| < \varepsilon.$$

Constrained problems

Inequality constrained DC optimisation problem:

$$\begin{aligned} &\min\ f_0(x) = g_0(x) - h_0(x) \\ &\text{subject to}\ f_i(x) = g_i(x) - h_i(x) \leq 0, \quad i \in I = \{1, \ldots, m\}. \end{aligned}$$

Constrained problems

Inequality constrained DC optimisation problem:

$$\begin{aligned} &\min \; f_0(x) = g_0(x) - h_0(x) \\ &\text{subject to} \;\; f_i(x) = g_i(x) - h_i(x) \le 0, \quad i \in I = \{1, \ldots, m\}. \end{aligned}$$

Similar approach:

$$\begin{aligned} &\min \; g_0(x) - \langle v_0, x - x_n \rangle \\ &\text{subject to} \quad g_i(x) - h_i(x_n) - \langle v_i, x - x_n \rangle \le 0, \quad i \in I. \end{aligned}$$

References

1. T. Lipp, S. Boyd, Variations and extension of the convex-concave procedure, Optimization and Engineering. 17, 263–287, 2016.
2. H.A. Le Thi, T. Pham Dinh, DC programming and DCA: thirty years of developments. Math. Program. 169, 5–68, 2018.
3. W. van Ackooij, W. de Oliveira, Non-smooth DC-constrained optimization: constraint qualification and minimizing methodologies, Optim. Methods Softw. 34, 890–920, 2019.

Contents

Nonlinear semidefinite programming

Nonlinear semidefinite programming problems:

$$\min\ f(x) \quad \text{subject to} \quad F(x) \preceq 0,$$

where $F \colon \mathbb{R}^d \to \mathbb{S}^\ell$, and $\mathbb{S}^\ell$ is the space of symmetric matrices of order ℓ, and $A \preceq B$ iff $B - A$ is positive semidefinite.

[1] Michael Stingl, On the Solution of Nonlinear Semidefinite Programs by Augmented Lagrangian Methods, PhD Thesis.

Nonlinear semidefinite programming

Nonlinear semidefinite programming problems:

$$\min \ f(x) \quad \text{subject to} \quad F(x) \preceq 0,$$

where $F\colon \mathbb{R}^d \to \mathbb{S}^\ell$, and $\mathbb{S}^\ell$ is the space of symmetric matrices of order ℓ, and $A \preceq B$ iff $B - A$ is positive semidefinite.

Applications: material optimization, truss topology design, structural optimization with vibration and stability constraints, robust gain-scheduling and some decentralized control problems, problems of maximizing the minimal eigenfrequency of a given structure, optimal $\mathcal{H}_2/\mathcal{H}_\infty$-static output feedback problems, etc.[1]

[1]Michael Stingl, On the Solution of Nonlinear Semidefinite Programs by Augmented Lagrangian Methods, PhD Thesis.

Order-theoretic approach

Let a matrix-valued function $F \colon \mathbb{R}^d \to \mathbb{S}^\ell$ be given.

Definition 1

The function F is called *convex*, if for all $x_1, x_2 \in \mathbb{R}^d$ and $\alpha \in [0, 1]$ one has

$$F(\alpha x_1 + (1 - \alpha) x_2) \preceq \alpha F(x_1) + (1 - \alpha) F(x_2).$$

Order-theoretic approach

Let a matrix-valued function $F \colon \mathbb{R}^d \to \mathbb{S}^\ell$ be given.

Definition 1

The function F is called *convex*, if for all $x_1, x_2 \in \mathbb{R}^d$ and $\alpha \in [0, 1]$ one has

$$F(\alpha x_1 + (1 - \alpha)x_2) \preceq \alpha F(x_1) + (1 - \alpha)F(x_2).$$

Definition 2

The function F is called *DC* (*Difference-of-Convex*), if there exist convex functions $G, H \colon \mathbb{R}^d \to \mathbb{S}^\ell$ such that $F = G - H$. Any such representation of the function F (or, equivalently, any such pair of functions (G, H)) is called a *DC decomposition* of F.

Counterexample

Example 1

Let $d = 1$, $\ell = 2$, and

$$F(x) = \begin{pmatrix} 1 & x^2 \\ x^2 & 1 \end{pmatrix}.$$

Then for $x_1 = 1$ and $x_2 = -1$ one has

$$\alpha F(x_1) + (1-\alpha)F(x_2) - F(\alpha x_1 + (1-\alpha)x_2) = \begin{pmatrix} 0 & 1-(2\alpha-1)^2 \\ 1-(2\alpha-1)^2 & 0 \end{pmatrix}.$$

This matrix is *not* positive semidefinite for any $\alpha \in (0,1)$, which implies that the function F is nonconvex.

DC decomposition of C^2-functions

Theorem 1

Let F be twice continuously differentiable and suppose that there exists $M > 0$ such that $\|\nabla^2 F_{ij}(x)\|_F \le M$ for all $i, j \in \{1, \ldots, \ell\}$.

DC decomposition of C^2-functions

Theorem 1

Let F be twice continuously differentiable and suppose that there exists $M > 0$ such that $\|\nabla^2 F_{ij}(x)\|_F \le M$ for all $i, j \in \{1, \ldots, \ell\}$. Then the function F is DC and for any $\mu \ge \ell M$ the pair (G, H) with

$$G(x) = F(x) + \frac{\mu}{2}|x|^2 I_\ell, \quad H(x) = \frac{\mu}{2}|x|^2 I_\ell, \quad x \in \mathbb{R}^d,$$

is a DC decomposition of F.

DC decomposition of C^2-functions

Theorem 1

Let F be twice continuously differentiable and suppose that there exists $M > 0$ such that $\|\nabla^2 F_{ij}(x)\|_F \le M$ for all $i, j \in \{1, \ldots, \ell\}$. Then the function F is DC and for any $\mu \ge \ell M$ the pair (G, H) with

$$G(x) = F(x) + \frac{\mu}{2}|x|^2 I_\ell, \quad H(x) = \frac{\mu}{2}|x|^2 I_\ell, \quad x \in \mathbb{R}^d,$$

is a DC decomposition of F.

A matrix-valued function F is convex if and only if for any $z \in \mathbb{R}^\ell$ the real-valued function $x \mapsto \langle z, F(x)z \rangle$ is convex.

Componentwise approach

Definition 3

The function F is called *componentwise convex*, if each component $F_{ij}(\cdot)$, $i, j \in \{1, \ldots, \ell\}$, is convex.

Componentwise approach

Definition 3

The function F is called *componentwise convex*, if each component $F_{ij}(\cdot)$, $i, j \in \{1, \ldots, \ell\}$, is convex. The function F is called *componentwise DC*, if there exist componentwise convex functions $G, H \colon \mathbb{R}^d \to \mathbb{S}^\ell$ such that $F = G - H$. Any such representation of the function F (or, equivalently, any such pair of functions (G, H)) is called *a componentwise DC decomposition* of F.

An example

If F is convex, then for $z = e_i$ the function $F_{ii}(x) = \langle z, F(x)z \rangle$ is convex.

An example

If F is convex, then for $z = e_i$ the function $F_{ii}(x) = \langle z, F(x)z \rangle$ is convex.

Example 1

Let $d = 1$, $\ell = 2$, and

$$F(x) = \begin{pmatrix} 0.5x^2 & \sin x \\ \sin x & 0.5x^2 \end{pmatrix}.$$

An example

If F is convex, then for $z = e_i$ the function $F_{ii}(x) = \langle z, F(x)z \rangle$ is convex.

Example 1

Let $d = 1$, $\ell = 2$, and

$$F(x) = \begin{pmatrix} 0.5x^2 & \sin x \\ \sin x & 0.5x^2 \end{pmatrix}.$$

Then for all $z \in \mathbb{R}^2$ and $x \in \mathbb{R}$ one has

$$\frac{d^2}{dx^2}\langle z, F(x)z \rangle = z_1^2 - 2(\sin x) z_1 z_2 + z_2^2 \geq z_1^2 - 2|z_1||z_2| + z_2^2$$
$$= (|z_1| - |z_2|)^2 \geq 0.$$

Thus, the function F is convex, despite the fact that non-diagonal elements of F are nonconvex.

What about non-diagonal elements?

What about non-diagonal elements?

Theorem 2

Let F be convex. Then for all $i, j \in \{1, \ldots, \ell\}$, $i \neq j$, the function F_{ij} is DC.

What about non-diagonal elements?

Theorem 2

Let F be convex. Then for all $i, j \in \{1, \ldots, \ell\}$, $i \neq j$, the function F_{ij} is DC.

Let $\ell = 2$. For $z = (1, 1)$ one has

$$\langle z, F(x)z \rangle = F_{11}(x) + 2F_{12}(x) + F_{22}(x),$$

which implies

$$F_{12}(x) = F_{21}(x) = \langle z, F(x)z \rangle - \Big(F_{11}(x) + F_{22}(x) \Big).$$

Corollaries

Corollary 2

Let F be convex. Then F is Lipschitz continuous on bounded sets.

Corollaries

Corollary 2

Let F be convex. Then F is Lipschitz continuous on bounded sets.

Corollary 3 (Aleksandrov-Busemann-Feller theorem for matrix-valued functions)

Let F be convex. Then F is twice differentiable almost everywhere.

Corollaries

Corollary 2

Let F be convex. Then F is Lipschitz continuous on bounded sets.

Corollary 3 (Aleksandrov-Busemann-Feller theorem for matrix-valued functions)

Let F be convex. Then F is twice differentiable almost everywhere.

Corollary 4

Any matrix-valued DC function F is componentwise DC.

Contents

Quadratic/Bilinear constraints

Suppose that

$$F(x) = C + \sum_{i=1}^{d} x_i B_i + \sum_{i,j=1}^{d} x_i x_j A_{ij} \qquad (1)$$

In particular, one can suppose that F is bilinear/biaffine, that is,

$$F(x, y) = A_{00} + \sum_{i=1}^{d} x_i A_{i0} + \sum_{j=1}^{m} y_j A_{0j} + \sum_{i=1}^{d} \sum_{j=1}^{m} x_i y_j A_{ij}. \qquad (2)$$

Quadratic/Bilinear constraints

Suppose that

$$F(x) = C + \sum_{i=1}^{d} x_i B_i + \sum_{i,j=1}^{d} x_i x_j A_{ij} \qquad (1)$$

In particular, one can suppose that F is bilinear/biaffine, that is,

$$F(x, y) = A_{00} + \sum_{i=1}^{d} x_i A_{i0} + \sum_{j=1}^{m} y_j A_{0j} + \sum_{i=1}^{d} \sum_{j=1}^{m} x_i y_j A_{ij}. \qquad (2)$$

Examples: simultaneous stabilisation of single-input single-output linear systems by one fixed controller of a given order, robust gain-scheduling, maximizing the minimal eigenfrequency of a given structure, etc.

Quadratic/Bilinear constraints

For any $\mu \geq \ell \max\limits_{s,k \in \{1,\ldots,\ell\}} \sum_{i,j=1}^{d} [A_{ij}]_{sk}^2$ the pair

$$G(x) = C + \sum_{i=1}^{d} x_i B_i + \sum_{i,j=1}^{d} x_i x_j A_{ij} + \frac{\mu}{2}|x|^2 I_\ell, \quad H(x) = \frac{\mu}{2}|x|^2 I_\ell$$

is a DC decomposition of F.

Quadratic/Bilinear constraints

For any $\mu \geq \ell \max_{s,k \in \{1,\ldots,\ell\}} \sum_{i,j=1}^{d} [A_{ij}]_{sk}^2$ the pair

$$G(x) = C + \sum_{i=1}^{d} x_i B_i + \sum_{i,j=1}^{d} x_i x_j A_{ij} + \frac{\mu}{2}|x|^2 I_\ell, \quad H(x) = \frac{\mu}{2}|x|^2 I_\ell$$

is a DC decomposition of F. Let

$$A = \begin{pmatrix} A_{11} & \ldots & A_{1d} \\ \ldots & \ldots & \ldots \\ A_{d1} & \ldots & A_{dd} \end{pmatrix}$$

If a decomposition $A = A_+ + A_-$ is known, one can define

$$G(x) = C + \sum_{i=1}^{d} x_i B_i + \sum_{i,j=1}^{d} x_i x_j (A_+)_{ij}, \quad H(x) = -\sum_{i,j=1}^{d} x_i x_j (A_-)_{ij}.$$

Bilinear/Biaffine Matrix Constraints

Let

$$F(X_1, X_2, X_3) = \begin{bmatrix} X_1 & (A + BX_2 C)X_3 \\ X_3(A + BX_2 C)^T & X_3 \end{bmatrix} \preceq 0$$

for all $X_1, X_3 \in \mathbb{S}^{\ell}$, $X_2 \in \mathbb{R}^{m \times m}$.

Bilinear/Biaffine Matrix Constraints

Let

$$F(X_1, X_2, X_3) = \begin{bmatrix} X_1 & (A + BX_2C)X_3 \\ X_3(A + BX_2C)^T & X_3 \end{bmatrix} \preceq 0$$

for all $X_1, X_3 \in \mathbb{S}^\ell$, $X_2 \in \mathbb{R}^{m \times m}$. (**Examples**: optimal $\mathcal{H}_2/\mathcal{H}_\infty$-static output feedback problems.)

Bilinear/Biaffine Matrix Constraints

Let

$$F(X_1, X_2, X_3) = \begin{bmatrix} X_1 & (A + BX_2C)X_3 \\ X_3(A + BX_2C)^T & X_3 \end{bmatrix} \preceq 0$$

for all $X_1, X_3 \in \mathbb{S}^\ell$, $X_2 \in \mathbb{R}^{m \times m}$. (**Examples**: optimal $\mathcal{H}_2/\mathcal{H}_\infty$-static output feedback problems.) For any $\mu \geq \ell M$, where

$$M^2 = \max_{i \in \{1,\ldots,\ell\}} \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \sum_{k_3=1}^{\ell} \left(B_{ik_1} C_{k_2 k_3}\right)^2,$$

the pair

$$G(x) = F(x) + \frac{\mu}{2}\left(\|X_2\|_F^2 + \|X_3\|_F^2\right) I_{2\ell}, \quad H(x) = \frac{\mu}{2}\left(\|X_2\|_F^2 + \|X_3\|_F^2\right) I_{2\ell}$$

is a DC decomposition of F, where $\|X\|_F = \sqrt{\operatorname{Tr} X^2}$ is the Frobenius norm.

The Stiefel manifold/orthogonality constraint

Consider the equality constraint

$$X^T X = I_\ell, \tag{3}$$

which is known as the Stiefel manifold or orthogonality constraint appearing in many applications (e.g. multi-matrix principal component analysis).

[2] T. Lipp, S. Boyd, Variations and extension of the convex-concave procedure, Optimization and Engineering, 2016.

The Stiefel manifold/orthogonality constraint

Consider the equality constraint

$$X^T X = I_\ell, \tag{3}$$

which is known as the Stiefel manifold or orthogonality constraint appearing in many applications (e.g. multi-matrix principal component analysis).

We can rewrite the constraint as follows:

$$G(X) = X^T X - I_\ell \preceq 0, \quad H(X) = I_\ell - X^T X \preceq 0.$$

The functions G and $-H$ are convex.[2]

[2] T. Lipp, S. Boyd, Variations and extension of the convex-concave procedure, Optimization and Engineering, 2016.

Contents

Equivalent reformulation

Nonlinear semidefinite programming problem

$$\min\ f(x) \quad \text{subject to} \quad F(x) \preceq 0,$$

can be rewritten as

$$\min\ f(x) \quad \text{subject to} \quad \lambda_{\max}(F(x)) \leq 0.$$

Equivalent reformulation

Nonlinear semidefinite programming problem

$$\min\ f(x) \quad \text{subject to} \quad F(x) \preceq 0,$$

can be rewritten as

$$\min\ f(x) \quad \text{subject to} \quad \lambda_{\max}(F(x)) \leq 0.$$

Is $\lambda_{\max}(F(\cdot))$ DC, when F is componentwise DC?

Maximal eigenvalue

Theorem 3

Let F be componentwise DC and $F_{ij} = G_{ij} - H_{ij}$ be a DC decomposition of each component of F, $i, j \in \{1, \ldots, \ell\}$.

Maximal eigenvalue

Theorem 3

Let F be componentwise DC and $F_{ij} = G_{ij} - H_{ij}$ be a DC decomposition of each component of F, $i, j \in \{1, \ldots, \ell\}$. Then the function $\lambda_{\max}(F(\cdot))$ is DC and the pair (g, h) with

$$\begin{aligned} g(x) &= \max_{|v| \le 1} \sum_{i,j=1}^{\ell} \Big((v_i v_j + 1) G_{ij}(x) + (1 - v_i v_j) H_{ij}(x) \Big), \\ h(x) &= \sum_{i,j=1}^{\ell} \Big(G_{ij}(x) + H_{ij}(x) \Big) \end{aligned} \tag{4}$$

for all $x \in \mathbb{R}^d$ is a DC decomposition of the function $\lambda_{\max}(F(\cdot))$.

Maximal eigenvalue

Note that $g(x) = \lambda_{\max}(F(x)) + h(x)$.

Maximal eigenvalue

Note that $g(x) = \lambda_{\max}(F(x)) + h(x)$.

For any x one has

$$\partial g(x) = \operatorname{co}\Big\{ \sum_{i,j=1}^{\ell} \Big((v_i v_j + 1)\partial G_{ij}(x) + (1 - v_i v_j)\partial H_{ij}(x) \Big) \Bigm| v \in \mathcal{E}_{\max}(A) \colon |v| = 1 \Big\},$$

where $\mathcal{E}_{\max}(F(x))$ is the corresponding eigenspace.

Contents

Cone constrained DC optimisation

Consider the following problem:

$$\begin{aligned} &\min \quad f_0(x) = g_0(x) - h_0(x), \\ &\text{subject to} \quad F(x) = G(x) - H(x) \preceq_K 0, \quad x \in A. \end{aligned} \qquad (\mathcal{P})$$

Cone constrained DC optimisation

Consider the following problem:

$$\begin{aligned} &\min \ \ f_0(x) = g_0(x) - h_0(x), \\ &\text{subject to } \ F(x) = G(x) - H(x) \preceq_K 0, \quad x \in A. \end{aligned} \qquad (\mathcal{P})$$

Here K is a proper cone in a real Banach space Y, $\preceq_K$ is the partial order induced by the cone K, i.e. $x \preceq_K y$ iff $y - x \in K$,

Cone constrained DC optimisation

Consider the following problem:

$$\begin{aligned} &\min \quad f_0(x) = g_0(x) - h_0(x), \\ &\text{subject to} \quad F(x) = G(x) - H(x) \preceq_K 0, \quad x \in A. \end{aligned} \qquad (\mathcal{P})$$

Here K is a proper cone in a real Banach space Y, $\preceq_K$ is the partial order induced by the cone K, i.e. $x \preceq_K y$ iff $y - x \in K$, and F is DC with respect to this partial order, i.e. the functions $G, H \colon \mathbb{R}^d \to Y$ are convex with respect to the cone K (or K-convex):

$$G(\alpha x_1 + (1 - \alpha)x_2) \preceq_K \alpha G(x_1) + (1 - \alpha)G(x_2)$$

for all $\alpha \in [0, 1]$ and $x_1, x_2 \in \mathbb{R}^d$.

Optimality conditions

We suppose that H is Fréchet differentiable.

Optimality conditions

We suppose that H is Fréchet differentiable.

Theorem 4

Let x_* be a locally optimal solution of the problem $(\mathcal{P})$. Then for any $v \in \partial h_0(x_*)$ the point x_* is a globally optimal solutions of the convex problem:

$$\begin{aligned} &\min \quad g_0(x) - \langle v, x - x_* \rangle \\ &\text{subject to} \quad G(x) - H(x_*) - DH(x_*)(x - x_*) \preceq_K 0, \quad x \in A, \end{aligned} \tag{5}$$

where $DH(x_*)$ is the Fréchet derivative of H at x_*.

Optimality conditions

We suppose that H is Fréchet differentiable.

Theorem 4

Let x_* be a locally optimal solution of the problem $(\mathcal{P})$. Then for any $v \in \partial h_0(x_*)$ the point x_* is a globally optimal solutions of the convex problem:

$$\begin{aligned} &\min \quad g_0(x) - \langle v, x - x_* \rangle \\ &\text{subject to} \quad G(x) - H(x_*) - DH(x_*)(x - x_*) \preceq_K 0, \quad x \in A, \end{aligned} \tag{5}$$

where $DH(x_*)$ is the Fréchet derivative of H at x_*.

Points x_* that are optimal solutions of problem (5) are called *critical*.

Optimality conditions

Theorem 5

Suppose that the problem is smooth and

$$0 \in \operatorname{int} \Big\{ G(x) - H(x_*) - DH(x_*)(x - x_*) + K \Bigm| x \in A \Big\}.$$

Then x_* *is critical if and only if there exists a Lagrange multiplier* $\lambda_* \in K^*$ *such that* $\langle \lambda_*, F(x_*) \rangle = 0$ *and*

$$\langle D_x L(x_*, \lambda_*), x - x_* \rangle \geq 0 \quad \forall x \in A,$$

where $L(x, \lambda) = f_0(x) + \langle \lambda, F(x) \rangle$.

Optimality conditions

Theorem 5

Suppose that the problem is smooth and

$$0 \in \operatorname{int} \Big\{ G(x) - H(x_*) - DH(x_*)(x - x_*) + K \Bigm| x \in A \Big\}.$$

Then x_* is critical if and only if there exists a Lagrange multiplier $\lambda_* \in K^*$ such that $\langle \lambda_*, F(x_*) \rangle = 0$ and

$$\langle D_x L(x_*, \lambda_*), x - x_* \rangle \geq 0 \quad \forall x \in A,$$

where $L(x, \lambda) = f_0(x) + \langle \lambda, F(x) \rangle$.

If K has nonempty interior, then: $G(x) - H(x_*) - DH(x_*)(x - x_*) \in -\operatorname{int} K$ for some $x \in A$ (Slater's condition for problem (5)).

Contents

DCA for cone constrained problems

If x_0 is non-optimal, then one can find a "better" point by solving the convex problem

$$\begin{aligned} \min\quad & g_0(x) - \langle v, x - x_0 \rangle \\ \text{subject to}\quad & G(x) - H(x_0) - DH(x_0)(x - x_0) \preceq_K 0, \quad x \in A, \end{aligned}$$

for some $v \in \partial h_0(x_0)$. Interior point methods, augmented Lagrangian methods, etc. can be applied.[3]

[3] T. Lipp, S. Boyd, Variations and extension of the convex-concave procedure, Optimization and Engineering, 2016.

DCA for cone constrained problems

If x_0 is non-optimal, then one can find a "better" point by solving the convex problem

$$\begin{aligned} &\min \quad g_0(x) - \langle v, x - x_0 \rangle \\ &\text{subject to} \quad G(x) - H(x_0) - DH(x_0)(x - x_0) \preceq_K 0, \quad x \in A, \end{aligned}$$

for some $v \in \partial h_0(x_0)$. Interior point methods, augmented Lagrangian methods, etc. can be applied.[3]

If x_1 is a solution, then x_1 is feasible, and

$$\begin{aligned} f_0(x_1) = g_0(x_1) - h_0(x_1) &\leq g_0(x_1) - h_0(x_0) - \langle v, x_1 - x_0 \rangle \\ &\leq g_0(x_0) - h_0(x_0) = f_0(x_0), \end{aligned}$$

i.e. $f_0(x_1) \leq f_0(x_0)$. If x_0 is not critical, then $f_0(x_1) < f_0(x_0)$.

[3] T. Lipp, S. Boyd, Variations and extension of the convex-concave procedure, Optimization and Engineering, 2016.

DCA for cone constrained problems

Algorithm 3: DC Algorithm/The Convex-Concave Procedure (CCP).

Step 1. Choose a feasible initial point x_0 and set $n := 0$.
Step 2. Compute $v_n \in \partial h_0(x_n)$ and $DH(x_n)$.
Step 3. Set the value of x_{n+1} to a solution of the convex problem

$$\begin{aligned} &\min \ \ g_0(x) - \langle v_n, x - x_n \rangle \\ &\text{subject to} \ \ G(x) - H(x_n) - DH(x_n)(x - x_n) \preceq_K 0, \quad x \in A. \end{aligned}$$

If $x_{n+1} = x_n$, **Stop**. Otherwise, put $n := n + 1$ and go to **Step 2**.

DCA for cone constrained problems

Theorem 5

Let f_0 be bounded below on the feasible region. Then the following statements hold true:

DCA for cone constrained problems

Theorem 5

Let f_0 be bounded below on the feasible region. Then the following statements hold true:

1. *the sequence $\{x_n\}$ is feasible for the problem $(\mathcal{P})$;*

DCA for cone constrained problems

Theorem 5

Let f_0 be bounded below on the feasible region. Then the following statements hold true:

1. *the sequence $\{x_n\}$ is feasible for the problem $(\mathcal{P})$;*
2. *for any $n \in \mathbb{N} \cup \{0\}$ either x_n is critical and the process terminates at step n or $f_0(x_{n+1}) < f_0(x_n)$; moreover, if the algorithm does not terminate, then the sequence $\{f_0(x_n)\}$ converges;*

DCA for cone constrained problems

Theorem 5

Let f_0 be bounded below on the feasible region. Then the following statements hold true:

1. *the sequence $\{x_n\}$ is feasible for the problem $(\mathcal{P})$;*
2. *for any $n \in \mathbb{N} \cup \{0\}$ either x_n is critical and the process terminates at step n or $f_0(x_{n+1}) < f_0(x_n)$; moreover, if the algorithm does not terminate, then the sequence $\{f_0(x_n)\}$ converges;*
3. *if h_0 is strongly convex with constant $\mu > 0$, then*
$$f_0(x_{n+1}) \le f_0(x_n) - \frac{\mu}{2}|x_{n+1} - x_n|^2; \tag{6}$$
4. *if x_* is a limit point of the sequence $\{x_n\}$ such that*
$$0 \in \operatorname{int}\Big\{ G(x) - H(x_*) - DH(x_*)(x - x_*) + K \,\Big|\, x \in A \Big\},$$
then x_* is critical.

DCA2/Penalty Convex-Concave Procedure

Algorithm 4: DCA2/Penalty CCP.

Step 1. Choose an initial point $x_0 \in A$, penalty parameter $t_0 \succ_{K^*} 0$, the maximal norm of the penalty parameter $\tau_{\max} > 0$, $\mu > 1$, infeasibility tolerance $\varkappa > 0$, and set $n := 0$.
Step 2. Compute $v_n \in \partial h_0(x_n)$ and $DH(x_n)$.
Step 3. Set the value of x_{n+1} to a solution of the convex problem

$$\min_{(x,s)} \; g_0(x) - \langle v, x - x_n \rangle + \langle t_n, s \rangle$$

$$\text{subject to} \;\; G(x) - H(x_n) - DH(x_n)(x - x_n) \preceq_K s, \quad s \succeq_K 0, \quad x \in A.$$

If $x_{n+1} = x_n$, **Stop**.
Step 4. Define

$$t_{n+1} = \begin{cases} \mu t_n, & \text{if } \|s_{n+1}\| \geq \varkappa \text{ and } \mu \|t_n\| \leq \tau_{\max}, \\ t_n, & \text{otherwise,} \end{cases}$$

DCA2/Penalty CCP

Let the problem have the form

$$\min \; f_0(x) \quad \text{subject to} \quad f_i(x) = g_i(x) - h_i(x) \le 0, \quad i \in \{1, \ldots, m\}.$$

One can define $K = \mathbb{R}^m_+$.

DCA2/Penalty CCP

Let the problem have the form

$$\min\ f_0(x) \quad \text{subject to} \quad f_i(x) = g_i(x) - h_i(x) \le 0, \quad i \in \{1, \ldots, m\}.$$

One can define $K = \mathbb{R}^m_+$. Then $t_0 = (t_0^{(1)}, \ldots, t_0^{(m)})$, $t_0^{(i)} > 0$, and penalty subproblem can be rewritten as

$$\min_{x \in A} g_0(x) - \langle v, x - x_n \rangle + \sum_{i=1}^{m} t_n^{(i)} \max\{0, g_i(x) - h_i(x) - \langle v, x - x_n \rangle\}.$$

DCA2/Penalty CCP

Let the problem have the form

$$\min\ f_0(x) \quad \text{subject to} \quad f_i(x) = g_i(x) - h_i(x) \le 0, \quad i \in \{1, \ldots, m\}.$$

One can define $K = \mathbb{R}^m_+$. Then $t_0 = (t_0^{(1)}, \ldots, t_0^{(m)})$, $t_0^{(i)} > 0$, and penalty subproblem can be rewritten as

$$\min_{x \in A}\ g_0(x) - \langle v, x - x_n \rangle + \sum_{i=1}^{m} t_n^{(i)} \max\{0, g_i(x) - h_i(x) - \langle v, x - x_n \rangle\}.$$

DCA2 is, in essence, an application of DCA to the weighted ℓ_1 penalty function

$$\Phi_t(x) = f_0(x) + \sum_{i=1}^{m} t^{(i)} \max\{f_i(x), 0\}.$$

DCA2/Penalty CCP

The constraint $F(x) \preceq_K 0$ can be rewritten as $F(x) \in -K$.

DCA2/Penalty CCP

The constraint $F(x) \preceq_K 0$ can be rewritten as $F(x) \in -K$.

Lemma 1

Let Y be finite dimensional, the cone K be generating (i.e. $K - K = Y$), and the penalty function $\Phi_c(\cdot) = f_0(\cdot) + c \operatorname{dist}(F(\cdot), -K)$ be coercive on A for some $c > 0$. Then the iterations of DCA2/Penalty CCP are correctly defined, provided $\|t_0\|$ is sufficiently large.

DCA2/Penalty CCP

Definition 4

A point $x_* \in A$ is said to be *a generalized critical point* for vector $t \succ_{K^*} 0$, if there exist $v_* \in \partial h_0(x_*)$ and $s_* \succeq_K 0$ such that the pair (x_*, s_*) is a globally optimal solution of the problem

$$\begin{aligned} \min_{(x,s)} \quad & g_0(x) - \langle v_*, x - x_* \rangle + \langle t, s \rangle \\ \text{s.t.} \quad & G(x) - H(x_*) - DH(x_*)(x - x_*) \preceq_K s, \quad s \succeq_K 0, \quad x \in A. \end{aligned} \tag{7}$$

DCA2/Penalty CCP

Definition 4

A point $x_* \in A$ is said to be *a generalized critical point* for vector $t \succ_{K^*} 0$, if there exist $v_* \in \partial h_0(x_*)$ and $s_* \succeq_K 0$ such that the pair (x_*, s_*) is a globally optimal solution of the problem

$$\begin{aligned} &\min_{(x,s)} \; g_0(x) - \langle v_*, x - x_* \rangle + \langle t, s \rangle \\ &\text{s.t.} \; G(x) - H(x_*) - DH(x_*)(x - x_*) \preceq_K s, \quad s \succeq_K 0, \quad x \in A. \end{aligned} \tag{7}$$

Under some additional assumptions x_* is a generalized critical point if and only if $0 \in \partial \Phi_t(x_*)$, where ∂ is the Dini subdifferential.

DCA2/Penalty CCP

Lemma 2

Let a sequence $\{(x_n, s_n)\}$ be generated by Algorithm 4. Then

$$f_0(x_{n+1}) + \langle t_n, s_{n+1} \rangle \le f_0(x_n) + \langle t_n, s_n \rangle, \quad \forall n \in \mathbb{N}.$$

and this inequality is strict, if x_n is not a generalized critical point for t_n.

DCA2/Penalty CCP

Lemma 2

Let a sequence $\{(x_n, s_n)\}$ be generated by Algorithm 4. Then

$$f_0(x_{n+1}) + \langle t_n, s_{n+1} \rangle \le f_0(x_n) + \langle t_n, s_n \rangle, \quad \forall n \in \mathbb{N}.$$

and this inequality is strict, if x_n is not a generalized critical point for t_n.

The inequalities

$$\Big| f_0(x_{n+1}) + \langle t_n, s_{n+1} \rangle - f_0(x_n) + \langle t_n, s_n \rangle \Big| < \varepsilon, \quad \|s_n\| < \varepsilon_{feas}$$

can be used as a stopping criterion.

DCA2/Penalty CCP

Theorem 6

Let Y be finite dimensional, K be generating, and the penalty function $\Phi_c(x) = f_0(x) + c\,\mathrm{dist}(F(x), -K)$ be bounded below on A for

$$c = \min\{\langle t_0, s\rangle \mid s \in K,\ \|s\| = 1\} > 0.$$

DCA2/Penalty CCP

Theorem 6

Let Y be finite dimensional, K be generating, and the penalty function $\Phi_c(x) = f_0(x) + c\,\mathrm{dist}(F(x), -K)$ be bounded below on A for

$$c = \min\{\langle t_0, s\rangle \mid s \in K,\ \|s\| = 1\} > 0.$$

Then all limits points of the sequence $\{x_n\}$ generated by Algorithm 4 are generalized critical points for $t_* = \lim t_n$.

DCA2/Penalty CCP

Let $\varkappa = 0$ and $\tau_{\max} = +\infty$.

Let $\varkappa = 0$ and $\tau_{\max} = +\infty$.

Theorem 7

Let K be finite dimensional and $\Phi_c(\cdot) = f_0(\cdot) + c \operatorname{dist}(F(\cdot), -K)$ be coercive on A for some $c \geq 0$. Suppose also that $\{x_n\}$ converges to a point x_* satisfying the following constraint qualification:

$$0 \in \operatorname{int} \Big\{ G(x) - H(x_*) - DH(x_*)(x - x_*) + K \ \Big| \ x \in A \Big\}. \tag{8}$$

DCA2/Penalty CCP

Let $\varkappa = 0$ and $\tau_{\max} = +\infty$.

Theorem 7

Let K be finite dimensional and $\Phi_c(\cdot) = f_0(\cdot) + c\,\mathrm{dist}(F(\cdot), -K)$ be coercive on A for some $c \geq 0$. Suppose also that $\{x_n\}$ converges to a point x_* satisfying the following constraint qualification:

$$0 \in \operatorname{int}\Big\{ G(x) - H(x_*) - DH(x_*)(x - x_*) + K \;\Big|\; x \in A \Big\}. \tag{8}$$

Then the sequence $\{t_n\}$ is bounded, there exists $m \in \mathbb{N}$ such that for all $n \geq m$ the point x_n is feasible for the problem $(\mathcal{P})$, and the point x_* is feasible and critical for the problem $(\mathcal{P})$.

The End!

Thank you!