Optimization Problems with Geometric Constraints: Asymptotic Stationarity and an Augmented Lagrangian Method

Patrick Mehlitz

Brandenburgische Technische Universität Cottbus - Senftenberg

Joint ongoing work with
Matůs Benko, Christian Kanzow, Alexander Y. Kruger, and Gerd Wachsmuth

November 01, 2021

Brandenburgische Technische Universität Cottbus - Senftenberg

1. Introduction
2. Asymptotic stationarity and regularity
3. A safeguarded augmented Lagrangian method
4. Numerical results
5. Related work

Problem setting

We consider the mathematical program

$$\begin{aligned} f(x) &\to \min \\ G(x) &\in K \\ x &\in C \end{aligned} \tag{MPGC}$$

for continuously differentiable data functions $f\colon \mathbb{X} \to \mathbb{R}$ and $G\colon \mathbb{X} \to \mathbb{Y}$ where $\mathbb{X}$ and $\mathbb{Y}$ are Euclidean spaces, $K \subset \mathbb{Y}$ is convex and closed, while $C \subset \mathbb{X}$ is closed and, potentially, of challenging *geometric* structure.

We consider the mathematical program

$$\begin{aligned} f(x) &\to \min \\ G(x) &\in K \qquad\qquad \text{(MPGC)} \\ x &\in C \end{aligned}$$

for continuously differentiable data functions $f\colon \mathbb{X} \to \mathbb{R}$ and $G\colon \mathbb{X} \to \mathbb{Y}$ where $\mathbb{X}$ and $\mathbb{Y}$ are Euclidean spaces, $K \subset \mathbb{Y}$ is convex and closed, while $C \subset \mathbb{X}$ is closed and, potentially, of challenging *geometric* structure.

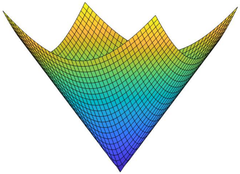
We call (MPGC) a *mathematical program with geometric constraints*. The feasible set of (MPGC) will be denoted by $\mathcal{F}$.

The set C is a closed, convex cone.

The set C is a closed, convex cone.

Some examples:

- semidefinite programming, i.e., $\mathbb{X} := \mathbb{R}^{n\times n}_{\text{sym}}$ and $C := \{X \in \mathbb{R}^{n\times n}_{\text{sym}} \,|\, X \succeq 0\}$: eigenvalue optimization, matrix inequality constraints (communication theory, experimental design)
- second-order cone programming, i.e., $\mathbb{X} := \mathbb{R}^n \times \mathbb{R}$ and $C := \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \,|\, \|x\| \leq t\}$: reformulations of probabilistic or robustified constraints



A second-order cone in $\mathbb{R}^2 \times \mathbb{R}$.

Disjunctive programs

The set C is the union of finitely many poyhedral sets (so-called *disjunctive*).

The set C is the union of finitely many poyhedral sets (so-called *disjunctive*).

Some examples:

- mathematical programs with complementarity constraints (MPCCs)

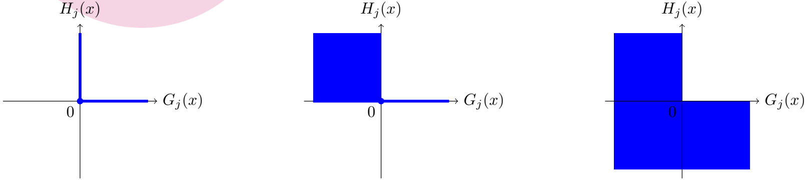
$$0 \leq G_j(x) \perp H_j(x) \geq 0 \quad j = 1, \ldots, q$$

- mathematical programs with vanishing constraints (MPVCs)

$$H_j(x) \geq 0 \quad G_j(x)\, H_j(x) \leq 0 \quad j = 1, \ldots, q,$$

- mathematical programs with or-constraints (MPOCs)

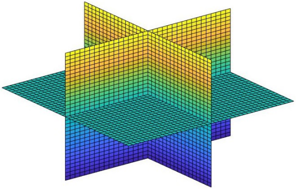
$$G_j(x) \leq 0 \quad \vee \quad H_j(x) \leq 0 \quad j = 1, \ldots, q.$$



For $\mathbb{X} := \mathbb{R}^n$ and a natural number $\kappa \in \{1, \ldots, n-1\}$, set $C := \{x \in \mathbb{R}^n \,|\, \|x\|_0 \leq \kappa\}$ where $\|\cdot\|_0 \colon \mathbb{R}^n \to \mathbb{R}$ counts the non-zero entries of a vector. Then (MPGC) amounts to a so-called *cardinality-constrained* optimization problem.

For $\mathbb{X} := \mathbb{R}^n$ and a natural number $\kappa \in \{1, \ldots, n-1\}$, set $C := \{x \in \mathbb{R}^n \,|\, \|x\|_0 \leq \kappa\}$ where $\|\cdot\|_0 \colon \mathbb{R}^n \to \mathbb{R}$ counts the non-zero entries of a vector. Then (MPGC) amounts to a so-called *cardinality-constrained* optimization problem.

- The search for *sparse* solutions of optimization problems is of essential importance in practical scenarios (data compression, portfolio optimization,...).
- The sparsity set C can be represented as the union of $\binom{n}{\kappa}$ κ-dimensional subspaces of $\mathbb{R}^n$ and, thus, is of challenging combinatorial structure.



A sparsity set in $\mathbb{R}^3$.

For $\mathbb{X} := \mathbb{R}^{m\times n}$ and a natural number $\kappa \in \{1,\ldots,\min(m,n)-1\}$, set $C := \{X \in \mathbb{R}^{m\times n} \mid \operatorname{rank} X \leq \kappa\}$. Then (MPGC) amounts to a so-called rank-constrained optimization problem.

For $\mathbb{X} := \mathbb{R}^{m\times n}$ and a natural number $\kappa \in \{1, \ldots, \min(m,n) - 1\}$, set $C := \{X \in \mathbb{R}^{m\times n} \mid \operatorname{rank} X \leq \kappa\}$. Then (MPGC) amounts to a so-called rank-constrained optimization problem.

Important applications for low-rank optimization can be found in machine learning, model order reduction, or matrix completion ("Netflix-Problem").

For $\mathbb{X} := \mathbb{R}^{m\times n}$ and a natural number $\kappa \in \{1,\ldots,\min(m,n)-1\}$, set $C := \{X \in \mathbb{R}^{m\times n} \mid \operatorname{rank} X \leq \kappa\}$. Then (MPGC) amounts to a so-called rank-constrained optimization problem.

Important applications for low-rank optimization can be found in machine learning, model order reduction, or matrix completion ("Netflix-Problem").

Reformulations of some problems from graph theory amount to rank-constrained matrix optimization problems. Exemplary, MAXCUT can be reformulated in $\mathbb{R}^{n\times n}_{\text{sym}}$ and then involves the constraint $\operatorname{rank} X \leq 1$.

1. Introduction

2. Asymptotic stationarity and regularity

3. A safeguarded augmented Lagrangian method

4. Numerical results

5. Related work

For a set-valued mapping $\Gamma\colon \mathbb{X} \rightrightarrows \mathbb{Y}$ and $\bar{x} \in \mathbb{X}$, we define

$$\limsup_{x\to\bar{x}} \Gamma(x) := \{y \in \mathbb{Y} \,|\, \exists \{(x_k, y_k)\}_{k\in\mathbb{N}} \subset \operatorname{gph}\Gamma,\, x_k \to \bar{x},\, y_k \to y\}\,,$$

the upper (or outer) limit of Γ at $\bar{x}$.

For a set-valued mapping $\Gamma\colon \mathbb{X} \rightrightarrows \mathbb{Y}$ and $\bar{x} \in \mathbb{X}$, we define

$$\limsup_{x \to \bar{x}} \Gamma(x) := \{y \in \mathbb{Y} \,|\, \exists \{(x_k, y_k)\}_{k \in \mathbb{N}} \subset \operatorname{gph} \Gamma,\, x_k \to \bar{x},\, y_k \to y\}\,,$$

the upper (or outer) limit of Γ at $\bar{x}$.

For a closed set $A \subset \mathbb{X}$ and $\bar{x} \in A$, we define

$$\mathcal{N}_A(\bar{x}) := \limsup_{x \to \bar{x}} \operatorname{cone}(x - \Pi_A(x)),$$

the limiting normal cone to A at $\bar{x}$. Here, $\Pi_A\colon \mathbb{X} \rightrightarrows \mathbb{X}$ is the (possibly set-valued) projection onto A. For $\tilde{x} \notin A$, we set $\mathcal{N}_A(\tilde{x}) := \varnothing$.

A feasible point $\bar{x} \in \mathcal{F}$ of (MPGC) is called M-stationary, whenever

$$-\nabla f(\bar{x}) \in G'(\bar{x})^* \mathcal{N}_K(G(\bar{x})) + \mathcal{N}_C(\bar{x})$$

holds. For NLPs, this corresponds to the standard KKT-conditions.

Brandenburgische Technische Universität Cottbus - Senftenberg

A feasible point $\bar{x} \in \mathcal{F}$ of (MPGC) is called M-stationary, whenever

$$-\nabla f(\bar{x}) \in G'(\bar{x})^* \mathcal{N}_K(G(\bar{x})) + \mathcal{N}_C(\bar{x})$$

holds. For NLPs, this corresponds to the standard KKT-conditions.

A local minimizer $\bar{x} \in \mathcal{F}$ of (MPGC) is an M-stationary point only in the presence of suitable constraint qualifications like GMFCQ:

$$-G'(\bar{x})^* \lambda \in \mathcal{N}_C(\bar{x}),\, \lambda \in \mathcal{N}_K(G(\bar{x})) \quad \Longrightarrow \quad \lambda = 0.$$

The latter corresponds to MFCQ for NLPs.

A feasible point $\bar{x} \in \mathcal{F}$ of (MPGC) is called M-stationary, whenever

$$-\nabla f(\bar{x}) \in G'(\bar{x})^* \mathcal{N}_K(G(\bar{x})) + \mathcal{N}_C(\bar{x})$$

holds. For NLPs, this corresponds to the standard KKT-conditions.

A local minimizer $\bar{x} \in \mathcal{F}$ of (MPGC) is an M-stationary point only in the presence of suitable constraint qualifications like GMFCQ:

$$-G'(\bar{x})^* \lambda \in \mathcal{N}_C(\bar{x}),\ \lambda \in \mathcal{N}_K(G(\bar{x})) \quad \Longrightarrow \quad \lambda = 0.$$

The latter corresponds to MFCQ for NLPs.

Is there a more general stationarity concept for (MPGC) which provides a necessary optimality condition in the absence of CQs and corresponds to the output of solution algorithms associated with (MPGC)? What do we need then in order to come up with M-stationarity?

We consider the following concept of *approximate* stationarity.

Definition

A feasible point $\bar{x} \in \mathcal{F}$ of (MPGC) is called *asymptotically stationary* whenever we find sequences $\{x_k\}_{k\in\mathbb{N}}, \{\varepsilon_k\}_{k\in\mathbb{N}} \subset \mathbb{X}$, $\{y_k\}_{k\in\mathbb{N}} \subset \mathbb{Y}$, and $\{\lambda_k\}_{k\in\mathbb{N}} \subset \mathbb{Y}$ satisfying $x_k \to \bar{x}$, $\varepsilon_k \to 0$, $y_k \to 0$, and

$$\forall k \in \mathbb{N}\colon \quad \varepsilon_k - \nabla f(x_k) - G'(x_k)^*\lambda_k \in \mathcal{N}_C(x_k), \quad \lambda_k \in \mathcal{N}_K(G(x_k) - y_k).$$

We consider the following concept of *approximate* stationarity.

Definition

A feasible point $\bar{x} \in \mathcal{F}$ of (MPGC) is called *asymptotically stationary* whenever we find sequences $\{x_k\}_{k\in\mathbb{N}}, \{\varepsilon_k\}_{k\in\mathbb{N}} \subset \mathbb{X}$, $\{y_k\}_{k\in\mathbb{N}} \subset \mathbb{Y}$, and $\{\lambda_k\}_{k\in\mathbb{N}} \subset \mathbb{Y}$ satisfying $x_k \to \bar{x}$, $\varepsilon_k \to 0$, $y_k \to 0$, and

$$\forall k \in \mathbb{N}: \quad \varepsilon_k - \nabla f(x_k) - G'(x_k)^*\lambda_k \in \mathcal{N}_C(x_k), \quad \lambda_k \in \mathcal{N}_K(G(x_k) - y_k).$$

- Each M-stationary point is asymptotically stationary.
- If $\{\lambda_k\}_{k\in\mathbb{N}}$ is bounded, taking the limit yields M-stationarity.
- If $\{\lambda_k\}_{k\in\mathbb{N}}$ is not bounded, taking the limit yields Fritz–John-type M-stationarity with leading multiplier 0.

Theorem

Let $\bar{x} \in \mathcal{F}$ be a local minimizer of (MPGC). *Then $\bar{x}$ is asymptotically stationary.*

Theorem

Let $\bar{x} \in \mathcal{F}$ be a local minimizer of (MPGC). *Then $\bar{x}$ is asymptotically stationary.*

For the proof, we investigate the penalized problem

$$\begin{aligned} f(x) + \tfrac{k}{2}\left(\operatorname{dist}(G(x) - y, K) + \|y\|^2\right) + \tfrac{1}{2}\|x - \bar{x}\|^2 \;\to\; \min_{x,y} \\ x \;\in\; C \cap \mathbb{B}_\delta(\bar{x}), \\ y \;\in\; \mathbb{B}_\delta(0) \end{aligned}$$

for each $k \in \mathbb{N}$ and sufficiently small $\delta > 0$. The associated sequence of global solutions can be used to construct the sequences appearing in the definition of asymptotic stationarity.

Let us define a set-valued mapping $\mathcal{M}\colon \mathbb{X}\times\mathbb{Y}\rightrightarrows\mathbb{X}$ by

$$\forall x\in\mathbb{X}\,\forall y\in\mathbb{Y}\colon\quad \mathcal{M}(x,y):=G'(x)^*\mathcal{N}_K(G(x)-y)+\mathcal{N}_C(x).$$

Let us define a set-valued mapping $\mathcal{M}\colon \mathbb{X}\times\mathbb{Y}\rightrightarrows\mathbb{X}$ by

$$\forall x\in\mathbb{X}\,\forall y\in\mathbb{Y}\colon\quad \mathcal{M}(x,y):=G'(x)^*\mathcal{N}_K(G(x)-y)+\mathcal{N}_C(x).$$

For a feasible point $\bar x\in\mathcal{F}$ of (MPGC), we find

$$\bar x \text{ M-stationary} \quad\Longleftrightarrow\quad -\nabla f(\bar x)\in\mathcal{M}(\bar x,0),$$

$$\bar x \text{ asymptotically stationary} \quad\Longleftrightarrow\quad -\nabla f(\bar x)\in\limsup_{x\to\bar x,\,y\to 0}\mathcal{M}(x,y).$$

Let us define a set-valued mapping $\mathcal{M}\colon \mathbb{X}\times\mathbb{Y}\rightrightarrows\mathbb{X}$ by

$$\forall x\in\mathbb{X}\,\forall y\in\mathbb{Y}\colon\quad \mathcal{M}(x,y):=G'(x)^*\mathcal{N}_K(G(x)-y)+\mathcal{N}_C(x).$$

For a feasible point $\bar{x}\in\mathcal{F}$ of (MPGC), we find

$$\begin{aligned} &\bar{x}\text{ M-stationary} &&\Longleftrightarrow\quad -\nabla f(\bar{x})\in\mathcal{M}(\bar{x},0),\\ &\bar{x}\text{ asymptotically stationary} &&\Longleftrightarrow\quad -\nabla f(\bar{x})\in\limsup_{x\to\bar{x},\,y\to0}\mathcal{M}(x,y).\end{aligned}$$

Definition

Let $\bar{x}\in X$ be feasible to (MPGC). Then $\bar{x}$ is said to be **asymptotically regular** whenever the subsequently stated condition holds:

$$\limsup_{x\to\bar{x},\,y\to0}\mathcal{M}(x,y)\subset\mathcal{M}(\bar{x},0).$$

Theorem

Let $\bar{x} \in \mathcal{F}$ be an asymptotically regular local minimizer of (MPGC). *Then $\bar{x}$ is M-stationary.*

Theorem

Let $\bar{x} \in \mathcal{F}$ be an asymptotically regular local minimizer of (MPGC)*. Then $\bar{x}$ is M-stationary.*

- Asymptotic regularity is much weaker than GMFCQ (i.e., metric regularity of feasibility map) or problem-tailored versions of RCPLD (in case $K := \mathbb{R}^m_- \times \{0\}^p$).
- Asymptotic regularity is independent of MSCQ (i.e., metric subregularity of feasibility map), ACQ, and GCQ.
- Asymptotic regularity is inherent whenever G is affine, K is polyhedral, and C is disjunctive.

Outline

1. Introduction

2. Asymptotic stationarity and regularity

3. A safeguarded augmented Lagrangian method

4. Numerical results

5. Related work

For some penalty parameter $\rho > 0$, let $\mathcal{L}_\rho\colon \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ denote the (partial) augmented Lagrangian function

$$\forall x \in \mathbb{X}\, \forall \lambda \in \mathbb{Y}\colon \quad \mathcal{L}_\rho(x,\lambda) := f(x) + \tfrac{\rho}{2}\operatorname{dist}^2\bigl(G(x) + \lambda/\rho, K\bigr).$$

For some penalty parameter $\rho > 0$, let $\mathcal{L}_\rho \colon \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ denote the (partial) augmented Lagrangian function

$$\forall x \in \mathbb{X}\, \forall \lambda \in \mathbb{Y}\colon \quad \mathcal{L}_\rho(x,\lambda) := f(x) + \tfrac{\rho}{2}\operatorname{dist}^2\bigl(G(x) + \lambda/\rho, K\bigr).$$

By convexity of K, $\mathcal{L}_\rho$ is a smooth function.

For some penalty parameter $\rho > 0$, let $\mathcal{L}_\rho\colon \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ denote the (partial) augmented Lagrangian function

$$\forall x \in \mathbb{X}\, \forall \lambda \in \mathbb{Y}\colon \quad \mathcal{L}_\rho(x,\lambda) := f(x) + \tfrac{\rho}{2}\operatorname{dist}^2\bigl(G(x) + \lambda/\rho, K\bigr).$$

By convexity of K, $\mathcal{L}_\rho$ is a smooth function.

Furthermore, we exploit the (partial) feasibility/complementarity measure $V_\rho\colon \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ given by

$$\forall x \in \mathbb{X}\, \forall \lambda \in \mathbb{Y}\colon \quad V_\rho(x,\lambda) := \bigl\|G(x) - \Pi_K\bigl(G(x) + \lambda/\rho\bigr)\bigr\|.$$

Algorithm 1 ALM for (MPGC)

procedure Safeguarded augmented Lagrangian method for (MPGC)

Input: Choose $\rho_0 > 0$, $\gamma > 1$, $\eta \in (0,1)$, some nonempty, bounded set $B \subset \mathbb{Y}$, $x_0 \in \mathbb{X}$. Set $k := 0$.

(S1) STOP whenever x_k satisfies a suitable termination criterion.

(S2) Choose $u_k \in B$ and solve $\min\{\mathcal{L}_{\rho_k}(x, u_k) \,|\, x \in C\}$ up to ε_{k+1}-M-stationarity, for small enough $\varepsilon_{k+1} \in \mathbb{X}$, i.e., find x_{k+1} such that

$$\varepsilon_{k+1} \in \nabla_x \mathcal{L}_{\rho_k}(x_{k+1}, u_k) + \mathcal{N}_C(x_{k+1}).$$

(S3) Set $\lambda_{k+1} := \rho_k \left(G(x_{k+1}) + u_k/\rho_k - \Pi_K \left(G(x_{k+1}) + u_k/\rho_k \right) \right)$.

(S4) If $k = 0$ or $V_{\rho_k}(x_{k+1}, u_k) \leq \eta\, V_{\rho_{k-1}}(x_k, u_{k-1})$, then $\rho_{k+1} := \rho_k$. Else, set $\rho_{k+1} := \gamma\, \rho_k$.

(S5) Set $k := k + 1$ and go to **(S1)**.

end procedure

Update of the multiplier (estimate):

- Replacing u_k by λ_k everywhere recovers the classical ALM.
- Safeguarding via the bounded set B enhances global convergence properties.
- In case where $\mathbb{Y}$ is equipped with a partial order relation, B is typically chosen as a (very large) box.
- One typically uses the multiplier estimate $u_k \in \Pi_B(\lambda_k)$.

Update of the multiplier (estimate):

- Replacing u_k by λ_k everywhere recovers the classical ALM.
- Safeguarding via the bounded set B enhances global convergence properties.
- In case where $\mathbb{Y}$ is equipped with a partial order relation, B is typically chosen as a (very large) box.
- One typically uses the multiplier estimate $u_k \in \Pi_B(\lambda_k)$.

Algorithm 1 is, at its core, a penalty method, so one can, generally, not force that accumulation points of the computed sequence are feasible to (MPGC). We use (approximate) feasibility w.r.t. the constraints $G(x) \in K$ as a termination criterion.

Theorem

Assume that the sequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$ *satisfies* $\varepsilon_k \to 0$. *Let* $\bar{x} \in \mathbb{X}$ *be an accumulation point of the sequence* $\{x_k\}_{k\in\mathbb{N}}$ *generated by Algorithm 1. Then the following assertions hold.*

- (i) *The point* $\bar{x}$ *is M-stationary for* $\min\{\operatorname{dist}^2(G(x), K) \,|\, x \in C\}$.
- (ii) *If* $\bar{x}$ *is feasible to* (MPGC)*, then it is asymptotically stationary.*
- (iii) *If* $\bar{x}$ *is an asymptotically regular feasible point of* (MPGC)*, then it is M-stationary.*

Theorem

Assume that the sequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$ *satisfies* $\varepsilon_k \to 0$. *Let* $\bar{x} \in \mathbb{X}$ *be an accumulation point of the sequence* $\{x_k\}_{k\in\mathbb{N}}$ *generated by Algorithm 1. Then the following assertions hold.*

(i) *The point* $\bar{x}$ *is M-stationary for* $\min\{\operatorname{dist}^2(G(x), K) \,|\, x \in C\}$.

(ii) *If* $\bar{x}$ *is feasible to* (MPGC), *then it is asymptotically stationary.*

(iii) *If* $\bar{x}$ *is an asymptotically regular feasible point of* (MPGC), *then it is M-stationary.*

For the solution of the ALM subproblems, we use a projected gradient method equipped with a Barzilai–Borwein-type nonmonotone line search.

- Projections onto C are often easy to compute (but not unique).
- This algorithm indeed computes approximate M-stationary of the subproblems.

1. Introduction

2. Asymptotic stationarity and regularity

3. A safeguarded augmented Lagrangian method

4. Numerical results

5. Related work

We consider a class of portfolio optimization problems given by

$$\min\{\tfrac{1}{2}x^\top\Sigma x \,|\, \mu^\top x \geq \rho,\ \mathsf{e}^\top x = 1,\ 0 \leq x \leq u,\ \|x\|_0 \leq \kappa\}$$

(covariance matrix $\Sigma \in \mathbb{R}^{n\times n}_{\mathsf{sym}}$, expected return $\mu \in \mathbb{R}^n$, minimum return ρ).
A test collection of problem instances based on random data has been set up by Frangioni/Gentile. We tackled all 30 test instances of dimension 200 with the three different values $\kappa \in \{5, 10, 20\}$ for each problem.

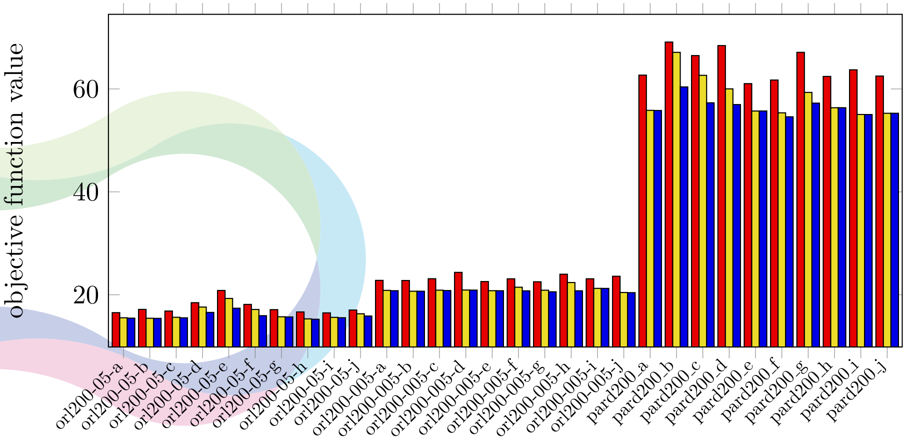
We consider a class of portfolio optimization problems given by

$$\min\{\tfrac{1}{2}x^\top \Sigma x \,|\, \mu^\top x \geq \rho,\ \mathsf{e}^\top x = 1,\ 0 \leq x \leq u,\ \|x\|_0 \leq \kappa\}$$

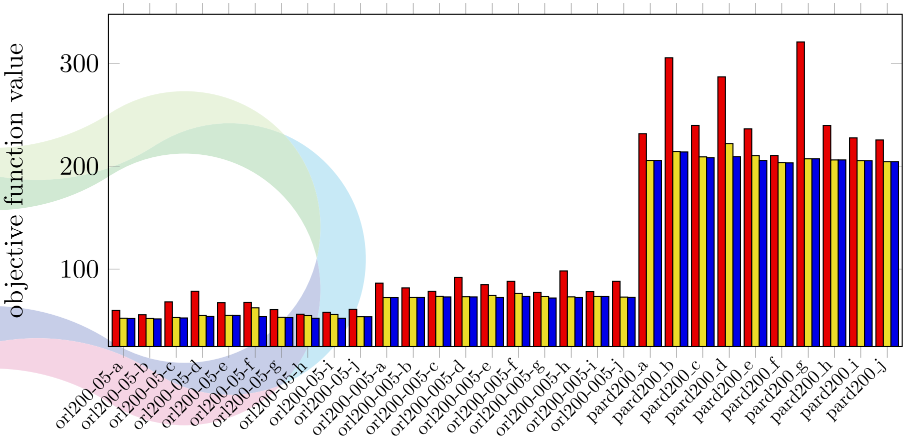
(covariance matrix $\Sigma \in \mathbb{R}^{n\times n}_{\mathsf{sym}}$, expected return $\mu \in \mathbb{R}^n$, minimum return ρ). A test collection of problem instances based on random data has been set up by Frangioni/Gentile. We tackled all 30 test instances of dimension 200 with the three different values $\kappa \in \{5, 10, 20\}$ for each problem.

For a numerical comparison, we compared Algorithm 1 (initialized at $x := 0$) with the performance of `CPLEX` and a boosted version of Algorithm 1 which finds a reasonable starting points via simple quadratic programming and exploits iterative reduction of κ, afterwards. For the implementation of Algorithm 1, we exploited projections onto

$$C := \{x \in [0, u] \,|\, \|x\|_0 \leq \kappa\}.$$



Optimal function values obtained by Algorithm 1 (red), boosted Algorithm 1 (yellow), and `CPLEX` (blue) with cardinality $\kappa := 20$.



Optimal function values obtained by Algorithm 1 (red), boosted Algorithm 1 (yellow), and `CPLEX` (blue) with cardinality $\kappa := 5$.

We fix an undirected, complete, weighted graph $\mathcal{G} = (V, E)$ with vertex set $V := \{1, \ldots, n\}$ and (symmetric) weight matrix $W \in \mathbb{R}^{n \times n}$.

We fix an undirected, complete, weighted graph $\mathcal{G} = (V, E)$ with vertex set $V := \{1, \ldots, n\}$ and (symmetric) weight matrix $W \in \mathbb{R}^{n \times n}$.

Setting $L := \operatorname{diag}(W\mathsf{e}) - W$, MAXCUT is equivalent to

$$\max\{\operatorname{trace}(LX) \,|\, \operatorname{diag} X = \mathsf{e},\ X \succeq 0,\ \operatorname{rank} X \leq 1\}$$

in $\mathbb{X} := \mathbb{R}^{n \times n}_{\mathsf{sym}}$. For the implementation of Algorithm 1, we used projections onto

$$C := \{X \in \mathbb{R}^{n \times n}_{\mathsf{sym}} \,|\, X \succeq 0,\ \operatorname{rank} X \leq 1\}.$$

We fix an undirected, complete, weighted graph $\mathcal{G} = (V, E)$ with vertex set $V := \{1, \ldots, n\}$ and (symmetric) weight matrix $W \in \mathbb{R}^{n\times n}$.

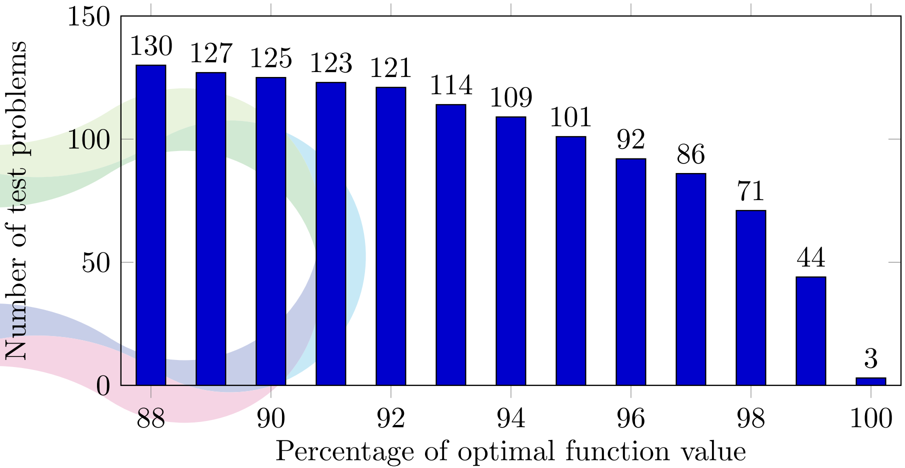
Setting $L := \operatorname{diag}(W\mathsf{e}) - W$, MAXCUT is equivalent to

$$\max\{\operatorname{trace}(LX) \,|\, \operatorname{diag} X = \mathsf{e},\ X \succeq 0,\ \operatorname{rank} X \leq 1\}$$

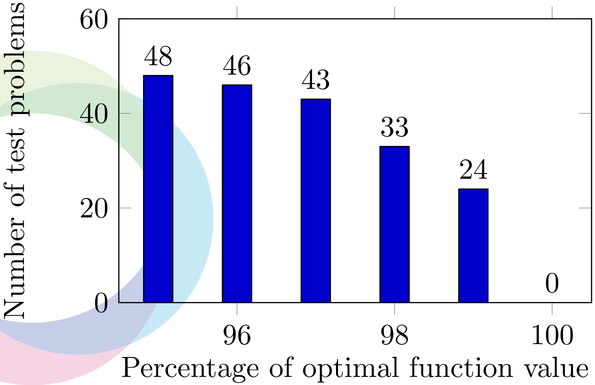
in $\mathbb{X} := \mathbb{R}^{n\times n}_{\mathsf{sym}}$. For the implementation of Algorithm 1, we used projections onto

$$C := \{X \in \mathbb{R}^{n\times n}_{\mathsf{sym}} \,|\, X \succeq 0,\ \operatorname{rank} X \leq 1\}.$$

We applied Algorithm 1 to the collections `rudy` (130 problems, $60 \leq n \leq 100$) and `ising` (48 problems, $100 \leq n \leq 400$) by Angelika Wiegele.



Results for the `rudy` collection.



Results for the `ising` collection.

1. Introduction

2. Asymptotic stationarity and regularity

3. A safeguarded augmented Lagrangian method

4. Numerical results

5. Related work

Bilevel optimization:

- algorithmic applications based on value function reformulation
- weak constraint qualifications

Bilevel optimization:

- algorithmic applications based on value function reformulation
- weak constraint qualifications

Optimization in infinite-dimensional Banach spaces:

- asymptotic stationarity under mild assumptions when C is convex as well
- constraint qualifications weaker than Robinson's CQ
- applies to safeguarded ALMs in Hilbert spaces

Bilevel optimization:

- algorithmic applications based on value function reformulation
- weak constraint qualifications

Optimization in infinite-dimensional Banach spaces:

- asymptotic stationarity under mild assumptions when C is convex as well
- constraint qualifications weaker than Robinson's CQ
- applies to safeguarded ALMs in Hilbert spaces

Nonsmooth optimization/variational calculus:

- generalization to nonsmooth (Lipschitzian) objective functions and generalized equation constraints possible
- asymptotic regularity serves as a qualification condition for the limiting variational calculus which is independent of metric subregularity, yields intersection rule for limiting normals and a chain rule for the coderivative calculus

Observations:

- local minimizers are either M-stationary or there is a critical direction u and an unbounded sequence of multipliers such that an asymptotic stationarity-type condition holds w.r.t. u and these multipliers
- asymptotic regularity is only necessary in critical directions

Observations:

- local minimizers are either M-stationary or there is a critical direction u and an unbounded sequence of multipliers such that an asymptotic stationarity-type condition holds w.r.t. u and these multipliers
- asymptotic regularity is only necessary in critical directions

Consequences:

- allows for the formulation of even weaker constraint qualifications and refined sufficient conditions in terms of problem data:
  * directional metric regularity
  * directional quasi-/pseudo-normality
  * directional polyhedrality
  * *horizon* coderivative criteria (handle unbounded multipliers), calculus challenging, results under metric pseudo (sub-)regularity of order 2
- qualification condition for directional limiting variational calculus

Asymptotic stationarity conditions in terms of Fréchet normals for problems of type

$$\min\{\varphi(x) \,|\, 0 \in \Phi(x)\}$$

where $\varphi\colon X \to \overline{\mathbb{R}}$ is lower semicontinuous and $\Phi\colon X \rightrightarrows Y$ is a closed-graph set-valued map between Asplund spaces; we need some *additional* lower semicontinuity w.r.t. the data

Asymptotic stationarity conditions in terms of Fréchet normals for problems of type

$$\min\{\varphi(x) \,|\, 0 \in \Phi(x)\}$$

where $\varphi\colon X \to \overline{\mathbb{R}}$ is lower semicontinuous and $\Phi\colon X \rightrightarrows Y$ is a closed-graph set-valued map between Asplund spaces; we need some *additional* lower semicontinuity w.r.t. the data

Applications:

- generalization of extremal principle (generalized set separation)
- qualification conditions for problems with non-Lipschitz objective functions (sparse portfolio selection, low rank matrix completion, edge-preserving image restoration)
- convergence analysis for multiplier-penalty-method for such problems
- optimality conditions for optimal control problems with sparsity-promoting term $u \mapsto \int_\Omega |u(\omega)|^p \,\mathrm{d}\omega$ for $p \in (0,1)$ on $L^2(\Omega)$

Thank you for your attention!