15.093 Optimization Methods

Lecture 24: Semidefinite Optimization

1 Outline

Slide 1

1. Minimizing Polynomials as an SDP
2. Linear Difference Equations and Stabilization
3. Barrier Algorithm for SDO

2 SDO formulation

2.1 Primal and dual

Slide 2

- $$\begin{array}{llll} (P): & \min & \boldsymbol{C} \bullet \boldsymbol{X} & \\ & \text{s.t.} & \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i & i = 1, \ldots, m \\ & & \boldsymbol{X} \succeq \boldsymbol{0} & \end{array}$$

- $$\begin{array}{lll} (D): & \max & \sum_{i=1}^{m} y_i b_i \\ & \text{s.t.} & \boldsymbol{C} - \sum_{i=1}^{m} y_i \boldsymbol{A}_i \succeq \boldsymbol{0} \end{array}$$

3 Minimizing Polynomials

3.1 Sum of squares

Slide 3

- A polynomial $f(x)$ is a **sum of squares** (SOS) if

  $$f(x) = \sum_j g_j^2(x)$$

  for some polynomials $g_j(x)$.
- A polynomial satisfies $f(x) \geq 0$ for all $x \in \mathcal{R}$ if and only if it is a sum of squares.
- **Not** true in more than one variable!

3.2 Proof

Slide 4

- ($\Leftarrow$) Obvious. If $f(x) = \sum_j g_j^2(x)$ then $f(x) \geq 0$.

- ($\Rightarrow$) Factorize $f(x) = C\prod_j (x-r_j)^{n_j} \prod_k (x-a_k+ib_k)^{m_k}(x-a_k-ib_k)^{m_k}$. Since $f(x)$ is nonnegative, then $C \geq 0$ and all the n_j are even. Then, $f(x) = f_1(x)^2 + f_2(x)^2$, where

$$
\begin{aligned}
f_1(x) &= C^{\frac{1}{2}} \prod_j (x-r_j)^{\frac{n_j}{2}} \prod_k (x-a_k)^{m_k} \\
f_2(x) &= C^{\frac{1}{2}} \prod_j (x-r_j)^{\frac{n_j}{2}} \prod_k b_k^{m_k}
\end{aligned}
$$

3.3 SOS and SDO

SLIDE 5

- Let $\boldsymbol{z}x = (1, x, x^2, \ldots, x^k)'$.
- $f(x) = \boldsymbol{z}(x)'\boldsymbol{Q}\boldsymbol{z}(x)$ is a sum of squares if and only if

$$f(x) = \boldsymbol{z}(x)'\boldsymbol{Q}\boldsymbol{z}(x),$$

  where $\boldsymbol{Q} \succeq \boldsymbol{0}$, i.e., $\boldsymbol{Q} = \boldsymbol{L'L}$.
- Then, $f(x) = \boldsymbol{z}(x)'\boldsymbol{L'L}(x) = ||L\boldsymbol{z}(x)||^2$.

3.4 Formulation

SLIDE 6

- Consider $\min f(x)$.
- Then, $f(x) \geq \gamma$ if and only if $f(x) - \gamma = \boldsymbol{z}x'\boldsymbol{Q}\boldsymbol{z}x$ with $\boldsymbol{Q} \succeq \boldsymbol{0}$. This implies linear constraints on γ and $\boldsymbol{Q}$.
- Reformulation

$$
\begin{aligned}
&\max \gamma \\
&\text{s.t.} \left\{ \begin{array}{rcl} f(x) - \gamma &=& \boldsymbol{z}(x)'\boldsymbol{Q}\boldsymbol{z}(x) \\ \boldsymbol{Q} &\succeq& \boldsymbol{0} \end{array} \right.
\end{aligned}
$$

3.5 Example

3.5.1 Reformulation

SLIDE 7

$$\min \ f(x) = 3 + 4x + 2x^2 + 2x^3 + x^4.$$

$$f(x) - \gamma = 3 - \gamma + 4x + 2x^2 + 2x^3 + x^4 = (1, x, x^2)'\boldsymbol{Q}(1, x, x^2).$$

$$
\begin{array}{rl}
\max & \gamma \\
\text{s.t.} & 3 - \gamma = q_{11} \\
& 4 = 2q_{12}, \quad 2 = 2q_{13} + q_{22} \\
& 2 = 2q_{23}, \quad 1 = q_{33} \\
& \boldsymbol{Q} = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \succeq \boldsymbol{0}
\end{array}
$$

Extensions to multiple dimensions.

4 Stability

Slide 8

- A linear difference equation

$$x(k+1) = \boldsymbol{A}x(k), \qquad x(0) = x_0$$

- $x(k)$ converges to zero iff $|\lambda_i(\boldsymbol{A})| < 1, i = 1, \dots n$
- Characterization:

$$|\lambda_i(\boldsymbol{A})| < 1 \quad \forall i \iff \exists \boldsymbol{P} \succ 0 \quad \boldsymbol{A}'\boldsymbol{P}\boldsymbol{A} - \boldsymbol{P} \prec 0$$

4.1 Proof

Slide 9

- ($\Longleftarrow$) Let $\boldsymbol{A}v = \lambda v$. Then,

$$0 > v'(\boldsymbol{A}'\boldsymbol{P}\boldsymbol{A} - \boldsymbol{P})v = (|\lambda|^2 - 1)\underbrace{v'\boldsymbol{P}v}_{>0},$$

  and therefore $|\lambda| < 1$

- ($\Longrightarrow$) Let $\boldsymbol{P} = \sum_{i=0}^{\infty} \boldsymbol{A}^{i'}\boldsymbol{Q}\boldsymbol{A}^i$, where $\boldsymbol{Q} \succ 0$. The sum converges by the eigenvalue assumption. Then,

$$\boldsymbol{A}'\boldsymbol{P}\boldsymbol{A} - \boldsymbol{P} = \sum_{i=1}^{\infty} \boldsymbol{A}^{i'}\boldsymbol{Q}\boldsymbol{A}^i - \sum_{i=0}^{\infty} \boldsymbol{A}^{i'}\boldsymbol{Q}\boldsymbol{A}^i = -\boldsymbol{Q} \prec 0$$

4.2 Stabilization

Slide 10

- Consider now the case where $\boldsymbol{A}$ is not stable, but we can change some elements, e.g., $\boldsymbol{A}(L) = \boldsymbol{A} + \boldsymbol{L}\boldsymbol{C}$, where $\boldsymbol{C}$ is a fixed matrix.
- Want to find an $\boldsymbol{L}$ such that $\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C}$ is stable.
- Use Schur complements to rewrite the condition:

$$(\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C})'\boldsymbol{P}(\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C}) - \boldsymbol{P} \prec 0, \qquad \boldsymbol{P} \succ 0$$
$$\Updownarrow$$
$$\begin{bmatrix} \boldsymbol{P} & (\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C})'\boldsymbol{P} \\ \boldsymbol{P}(\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C}) & \boldsymbol{P} \end{bmatrix} \succ 0$$

  Condition is nonlinear in $(\boldsymbol{P}, \boldsymbol{L})$

4.3 Changing variables

Slide 11

- Define a new variable $\boldsymbol{Y} := \boldsymbol{P}L$

$$\begin{bmatrix} \boldsymbol{P} & \boldsymbol{A}'\boldsymbol{P} + \boldsymbol{C}'\boldsymbol{Y}' \\ \boldsymbol{P}\boldsymbol{A} + \boldsymbol{Y}\boldsymbol{C} & \boldsymbol{P} \end{bmatrix} \succ 0$$

- This is linear in $(\boldsymbol{P}, \boldsymbol{Y})$.
- Solve using SDO, recover $\boldsymbol{L}$ via $\boldsymbol{L} = \boldsymbol{P}^{-1}\boldsymbol{Y}$

5 Primal Barrier Algorithm for SDO

SLIDE 12

- $\boldsymbol{X} \succeq \boldsymbol{0} \Leftrightarrow \lambda_1(\boldsymbol{X}) \geq 0, \ldots, \lambda_n(\boldsymbol{X}) \geq 0$
- Natural barrier to repel $\boldsymbol{X}$ from the boundary $\lambda_1(\boldsymbol{X}) > 0, \ldots, \lambda_n(\boldsymbol{X}) > 0$:

$$-\sum_{j=1}^{n} \log(\lambda_i(\boldsymbol{X})) =$$

$$-\log(\prod_{j=1}^{n} \lambda_i(\boldsymbol{X})) = -\log(\det(\boldsymbol{X}))$$

SLIDE 13

- Logarithmic barrier problem

$$\begin{array}{ll} \min & B_\mu(\boldsymbol{X}) = \boldsymbol{C} \bullet \boldsymbol{X} - \mu \log(\det(\boldsymbol{X})) \\ \text{s.t.} & \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i \quad, i = 1, \ldots, m, \\ & \boldsymbol{X} \succ \boldsymbol{0} \end{array}$$

- Derivative: $\nabla B_\mu(\boldsymbol{X}) = \boldsymbol{C} - \mu \boldsymbol{X}^{-1}$
  Follows from

$$\log\det(\boldsymbol{X} + \boldsymbol{H}) \approx \log\det(\boldsymbol{X}) + \text{trace}(\boldsymbol{X}^{-1}\boldsymbol{H}) + \cdots$$

- KKT conditions

$$\begin{array}{l} \boldsymbol{A}_i \bullet \boldsymbol{X} = \boldsymbol{b}_i \quad, i = 1, \ldots, m, \\ \boldsymbol{C} - \mu \boldsymbol{X}^{-1} = \sum_{i=1}^{m} y_i \boldsymbol{A}_i. \\ \boldsymbol{X} \succ \boldsymbol{0}, \end{array}$$

- Given μ, need to solve these nonlinear equations for X, C, y_i
- Apply Newton's method until we are "close" to the optimal
- Reduce value of μ, and iterate until the duality gap is small

5.1 Another interpretation

SLIDE 14

- Recall the optimality conditions:

$$\begin{array}{rl} \boldsymbol{A}_i \bullet \boldsymbol{X} & = \boldsymbol{b}_i \quad, i = 1, \ldots, m, \\ \sum_{i=1}^{m} y_i \boldsymbol{A}_i + \boldsymbol{S} & = \boldsymbol{C} \\ \boldsymbol{X}, \boldsymbol{S} & \succeq \boldsymbol{0}, \\ \boldsymbol{X}\,\boldsymbol{S} & = \boldsymbol{0} \end{array}$$

- Cannot solve directly. Rather, perturb the complementarity condition to $\boldsymbol{X}\,\boldsymbol{S} = \mu \boldsymbol{I}$.
- Now, unique solution for every $\mu > 0$ (the "central path")
- Solve using Newton, for decreasing values of μ.

6 Differences with LO

Slide 15

- Many different ways to linearize the nonlinear complementarity condition

$$\boldsymbol{X}\,\boldsymbol{S} = \mu \boldsymbol{I}$$

- Want to preserve symmetry of the iterates
- Several search directions.

7 Convergence

7.1 Stopping criterion

Slide 16

- The point $(\boldsymbol{X}, y_i)$ is feasible, and has duality gap:

$$\boldsymbol{C} \bullet \boldsymbol{X} - \sum_{i=1}^{m} y_i b_i = \mu \boldsymbol{X}^{-1} \bullet \boldsymbol{X} = n\mu$$

- Therefore, reducing μ always decreases the duality gap
- Barrier algorithm needs $O\left(\sqrt{n}\log\frac{\epsilon_0}{\epsilon}\right)$ iterations to reduce duality gap from ϵ_0 to ϵ

8 Conclusions

Slide 17

- SDO is a powerful modeling tool
- Barrier and primal-dual algorithms are very powerful
- Many good solvers available: SeDuMi, SDPT3, SDPA, etc.
- Pointers to literature and solvers:
`www-user.tu-chemnitz.de/~helmberg/semidef.html`

15.093J / 6.255J Optimization Methods
Fall 2009