2. SDP Relaxations for Quadratic Programming

- LQR with binary inputs
- Boolean optimization
- Primal and dual SDP relaxations
- Interpretations
- Examples
- S-procedure

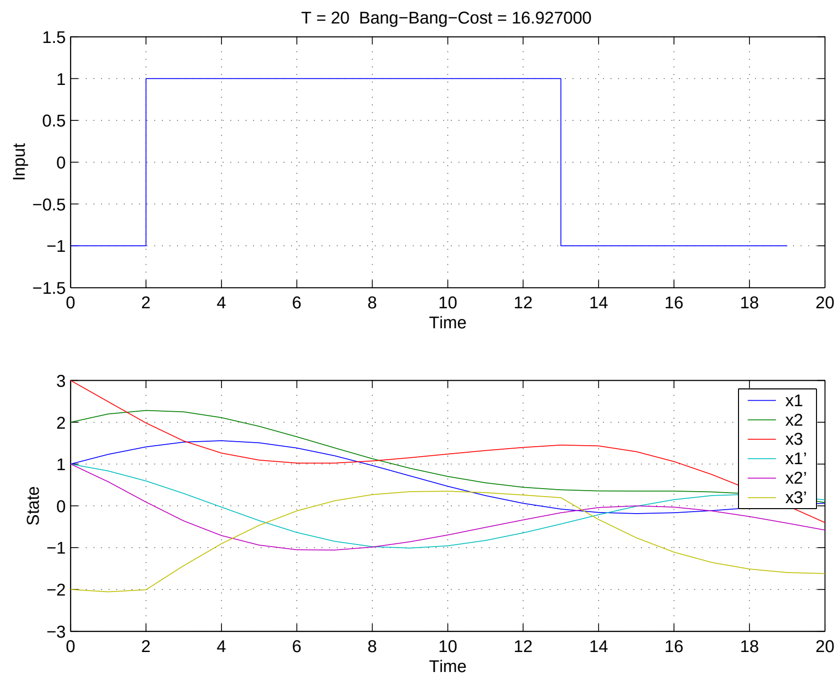
LQR with Binary Inputs

Consider the discrete-time LQR problem

$$\text{minimize} \quad \|y(t) - y_r(t)\|^2 \qquad \text{subject to} \quad \begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases}$$

where y_r is the reference output trajectory, and the input $u(t)$ is constrained by $|u(t)| = 1$ for all $t = 0, \ldots, N$.

An open-loop LQR-type problem, but with a *bang-bang* input.



LQR with Binary Inputs (continued)

The objective $\|y(t) - y_r(t)\|^2$ is a *quadratic* function of the input u:

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ y(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ CB & 0 & 0 & \dots & 0 \\ CAB & CB & 0 & \dots & 0 \\ \dots & \dots & \dots & \ddots & \vdots \\ CA^tB & CA^{t-1}B & \dots & CB & 0 \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \\ u(t) \end{bmatrix}$$

So the problem can be written as:

$$\begin{aligned} \text{minimize} \quad & \begin{bmatrix} u \\ 1 \end{bmatrix}^T \begin{bmatrix} Q & r \\ r^T & s \end{bmatrix} \begin{bmatrix} u \\ 1 \end{bmatrix} \\ \text{subject to} \quad & u_i \in \{+1, -1\} \text{ for all } i \end{aligned}$$

where Q, r, s are functions of the problem data.

This is a quadratic *boolean optimization* problem.

Boolean Minimization

A classic combinatorial problem:

$$\begin{array}{rl} \text{minimize} & x^T Q x \\ \text{subject to} & x_i \in \{-1, 1\} \end{array}$$

- Examples: MAX CUT, knapsack, LQR with binary inputs, etc.
- Can model the constraints with quadratic equations:

$$x_i^2 - 1 = 0 \quad \Longleftrightarrow \quad x_i \in \{-1, 1\}$$

- An exponential number of points. Cannot check them all!
- The problem is *NP-hard* (even if $Q \succeq 0$).

Despite the hardness of the problem, there are some very good approaches...

SDP Relaxations

$$\begin{array}{ll} \text{minimize} & x^T Q x \\ \text{subject to} & x_i^2 - 1 = 0 \end{array}$$

The Lagrangian function:

$$L(x, \lambda) = x^T Q x - \sum_{i=1}^{n} \lambda_i (x_i^2 - 1) = x^T (Q - \Lambda) x + \mathbf{trace}\, \Lambda$$

For the Lagrangian to be bounded below, we require $Q - \Lambda \succeq 0$.
The dual is therefore an SDP:

$$\begin{array}{ll} \text{maximize} & \mathbf{trace}\, \Lambda \\ \text{subject to} & Q - \Lambda \succeq 0 \end{array}$$

From this SDP we obtain a *primal-dual pair of SDP relaxations*

SDP Relaxations

$$\begin{array}{ll} \text{minimize} & x^T Q x \\ \text{subject to} & x_i^2 = 1 \end{array}$$

$$\begin{array}{ll} \text{minimize} & \mathbf{trace}\, QX \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1 \end{array} \qquad \begin{array}{ll} \text{maximize} & \mathbf{trace}\, \Lambda \\ \text{subject to} & Q \succeq \Lambda \\ & \Lambda \ \text{diagonal} \end{array}$$

- We derived them from Lagrangian and SDP duality
- But, these SDP relaxations arise in *many* other ways
- Well-known in combinatorial optimization, graph theory, etc.
- Several interpretations

SDP Relaxations: Dual Side

Gives an easy, "provable" *underestimator* of the objective function.

$$\begin{array}{ll} \text{maximize} & \textbf{trace}\,\Lambda \\ \text{subject to} & Q \succeq \Lambda \\ & \Lambda \ \text{diagonal} \end{array}$$

Directly provides a *lower bound* on the objective: for any feasible x:

$$x^T Q x \geq x^T \Lambda x = \sum_{i=1}^{n} \Lambda_{ii} x_i^2 = \textbf{trace}\,\Lambda$$

- The first inequality follows from $Q \succeq \Lambda$
- The second equation from Λ being diagonal
- The third, from $x_i \in \{+1, -1\}$

SDP Relaxations: Primal Side

The original problem is:

$$\begin{array}{rl} \text{minimize} & x^T Q x \\ \text{subject to} & x_i^2 = 1 \end{array}$$

Let $X := xx^T$. Then

$$x^T Q x = \mathbf{trace}\, Q x x^T = \mathbf{trace}\, QX$$

Therefore, $X \succeq 0$, has *rank one*, and $X_{ii} = x_i^2 = 1$.

Conversely, any matrix X with

$$X \succeq 0, \quad X_{ii} = 1, \quad \mathbf{rank}\, X = 1$$

necessarily has the form $X = xx^T$ for some ± 1 vector x.

Primal Side (continued)

Therefore, the original problem can be exactly rewritten as:

$$\begin{array}{rl} \text{minimize} & \mathbf{trace}\, QX \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1 \\ & \mathbf{rank}(X) = 1 \end{array}$$

Interpretation: "lift" to a higher dimensional space, from $\mathbb{R}^n$ to $\mathbb{S}^n$.
Dropping the (nonconvex) rank constraint, we obtain the relaxation.

If the solution X has rank 1, then we have solved the original problem.

Otherwise, *rounding schemes* to project solutions. In some cases, approximation guarantees (e.g. Goemans-Williamson for MAX CUT).

minimize	$\textbf{trace}\, QX$	maximize	$\textbf{trace}\, \Lambda$
subject to	$X \succeq 0$	subject to	$Q \succeq \Lambda$
	$X_{ii} = 1$		Λ diagonal

- Dual relaxations give *certified* bounds.
- Primal relaxations give information about possible *feasible* points.
- Both are solved *simultaneously* by primal-dual SDP solvers

Example

$$\begin{array}{rl} \text{minimize} & 2x_1x_2 + 4x_1x_3 + 6x_2x_3 \\ \text{subject to} & x_i^2 = 1 \end{array}$$

The associated matrix is $Q = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{bmatrix}$. The SDP solutions are:

$$X = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

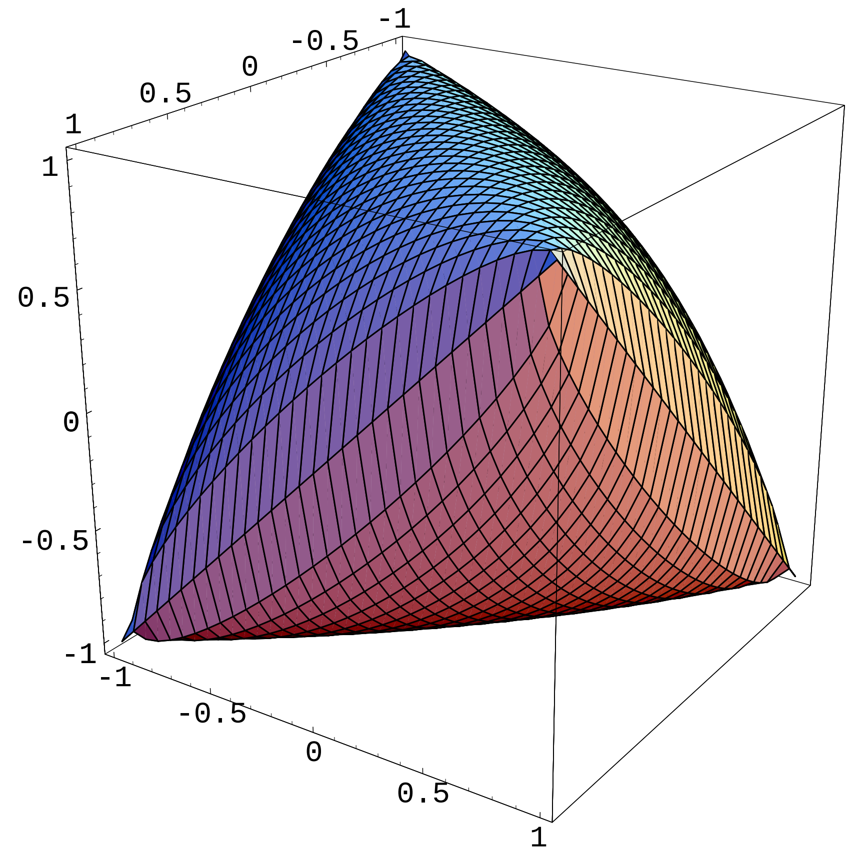
We have $X \succeq 0$, $X_{ii} = 1$, $Q - \Lambda \succeq 0$, and

$$\textbf{trace}\, QX = \textbf{trace}\, \Lambda = -8$$

Since X is rank 1, from $X = xx^T$ we recover the optimal $x = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}^T$, We can visualize this (in 3×3):

$$X = \begin{bmatrix} 1 & p_1 & p_2 \\ p_1 & 1 & p_3 \\ p_2 & p_3 & 1 \end{bmatrix} \succeq 0$$

in (p_1, p_2, p_3) space.



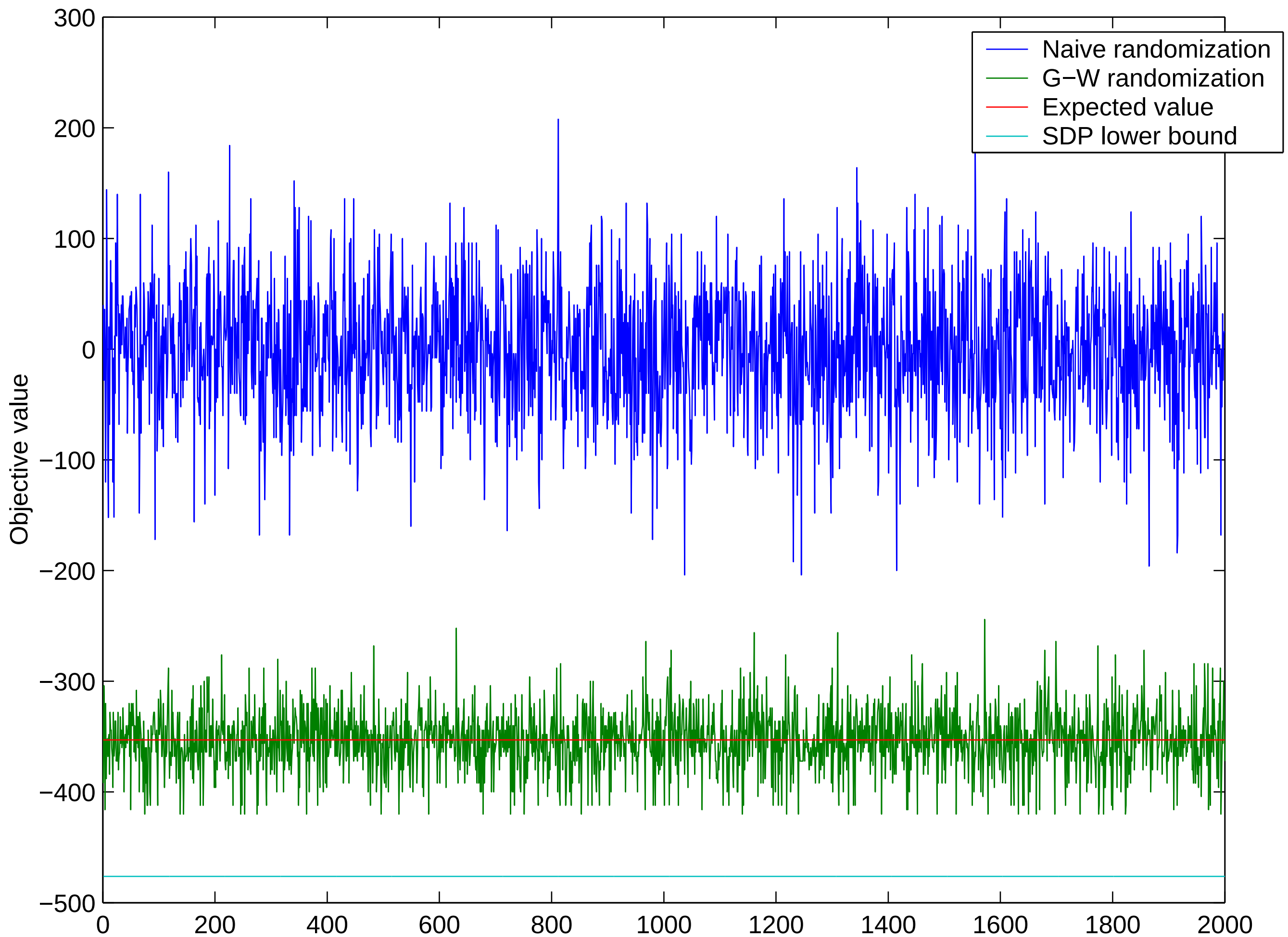
When optimizing the linear objective function

$$\textbf{trace}\, QX = 2p_1 + 4p_2 + 6p_3,$$

the optimal solution is at the "vertex" $(1, -1, -1)$.

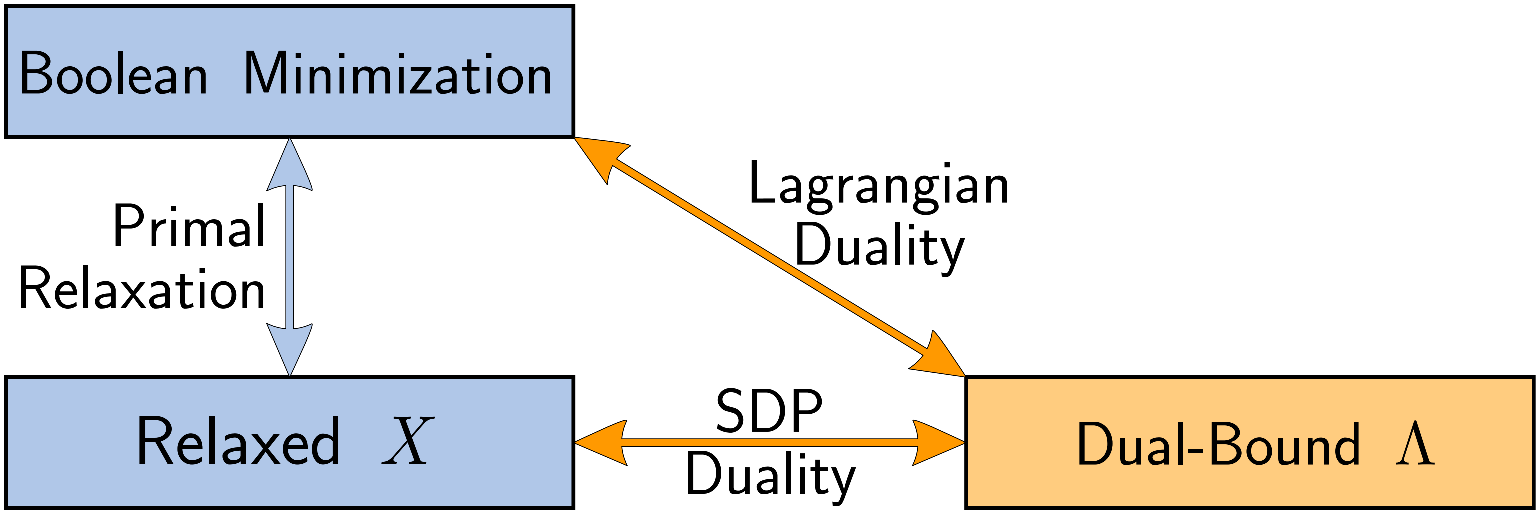
We can solve SDP relaxations of boolean QPs for problems of fairly large size (approx. 500 vars with interior point, 5000+ with special techniques).

Random example in 50 vars, computation time is around 1.5 sec.



SDP lower bound: -476.3198. G-W expected value: -352.9414.

A General Scheme



- The "relaxed" X suggests candidate points.
- The diagonal matrix Λ *certifies* a lower bound.

Ubiquitous scheme in optimization (convex hulls, fractional colorings, etc. . .)
We will learn systematic ways of constructing these, and more. . .

LQR with Binary Inputs (continued)

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} u \\ 1 \end{bmatrix}^T \begin{bmatrix} Q & r \\ r^T & s \end{bmatrix} \begin{bmatrix} u \\ 1 \end{bmatrix} \\ \text{subject to} & u_i \in \{+1, -1\} \text{ for all } i \end{array}$$

for some matrices (Q, r, s) function of the problem data (A, B, C, N).

An *SDP dual bound*:

$$\begin{array}{ll} \text{maximize} & \mathbf{trace}(\Lambda) + \mu \\ \text{subject to} & \begin{bmatrix} Q - \Lambda & r \\ r^T & s - \mu \end{bmatrix} \succeq 0, \qquad \Lambda \text{ diagonal} \end{array}$$

Let q^*, q_* be the optimal value of both problems. Then, $q^* \geq q_*$:

$$\begin{bmatrix} u \\ 1 \end{bmatrix}^T \begin{bmatrix} Q & r \\ r^T & s \end{bmatrix} \begin{bmatrix} u \\ 1 \end{bmatrix} \geq \begin{bmatrix} u \\ 1 \end{bmatrix}^T \begin{bmatrix} \Lambda & 0 \\ 0 & \mu \end{bmatrix} \begin{bmatrix} u \\ 1 \end{bmatrix} = \mathbf{trace}\,\Lambda + \mu$$

LQR with Binary Inputs (continued)

$$\begin{array}{ll} \text{maximize} & \mathbf{trace}(\Lambda) + \mu \\ \text{subject to} & \begin{bmatrix} Q - \Lambda & r \\ r^T & s - \mu \end{bmatrix} \succeq 0, \qquad \Lambda \text{ diagonal} \end{array}$$

Since $(\Lambda, \mu) = (0, 0)$ is always feasible, $q_* \geq 0$.

Furthermore, the bound is never worse than the LQR solution obtained by dropping the ± 1 constraint, since

$$\Lambda = 0, \quad \mu = s - r^T Q^{-1} r$$

is a feasible point.

Example:

N	LQR cost	SDP bound	Bang Bang
10	14.005	15.803	15.803
15	15.216	16.698	16.705
20	15.364	16.905	16.927

The S-procedure

A *sufficient* condition for the infeasibility of quadratic inequalities:

$$\{x \in \mathbb{R}^n \quad | \quad x^T A_i x \geq 0\}$$

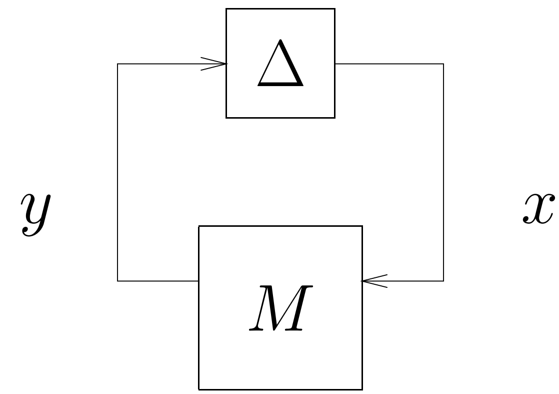
Again, a primal-dual pair of SDP relaxations:

$X \succeq 0$ $\mathbf{trace}\, X = 1$ $\mathbf{trace}\, A_i X \geq 0$	$\sum_i \lambda_i A_i \preceq -I$ $\lambda_i \geq 0$

The basis of many important results in control theory.

Structured Singular Value

- A central paradigm in robust control.
- μ is a measure of robustness: how big can a structured perturbation Δ be, without losing stability.



Do the loop equations admit nontrivial solutions?

$$y = Mx, \quad y_i^2 - x_i^2 \geq 0$$

Applying the standard SDP relaxation:

$$\sum_i d_i(y_i^2 - x_i^2) = x^T(M^T DM - D)x < 0, \qquad D = \mathbf{diag}(d_i),\, d_i \geq 0$$

We obtain the standard μ upper bound:

$$M^T DM - D \prec 0, \qquad D \quad \text{diagonal}, \quad D \succeq 0$$