Lecture 4: Nonlinear analysis of combinatorial problems

Yurii Nesterov, CORE/INMA (UCL)

September 12, 2012

Outline

Boolean quadratic problem

Let $Q = Q^T$ be an $(n \times n)$-matrix.

Maximization: find
$f^*(Q) \equiv \max\limits_x \{ \langle Qx, x \rangle : \; x_i = \pm 1, \; i = 1 \dots n \}$.

Minimization: find $f_*(Q) \equiv \min\limits_x \{ \langle Qx, x \rangle : \; x_i = \pm 1, \; i = 1 \dots n \}$.

Clearly $f^*(-Q) = -f_*(Q)$.

Trivial Properties

- Both problems are NP-hard.
- They can have up to 2^n local extremums.

Very often we are happy with approximate solutions

Simple bounds: Eigenvalues

Upper bound. For any $x \in R^n$ with $x_i = \pm 1$, we have $\|x\|^2 = n$. Therefore,

$$f^*(Q) \;\le\; \max_{\|x\|^2 = n} \langle Qx, x\rangle \;=\; n \cdot \lambda_{\max}(Q).$$

Lower bounds. 1. If $Q \succeq 0$, then

$$f^*(Q) = \max_{|x_i| \le 1} \langle Qx, x\rangle \;\ge\; \max_{\|x\|^2 = 1} \langle Qx, x\rangle \;=\; \lambda_{\max}(Q).$$

2. Consider random x with $\mathbf{Prob}\,(x_i = 1) = \mathbf{Prob}\,(x_i = -1) = \frac{1}{2}$. Then

$$f^*(Q) \;\ge\; E_x(\langle Qx, x\rangle) = \sum_{i,j=1}^{n} Q_{i,j} E_x(x_i x_j)$$

$$= \sum_{i=1}^{n} Q_{i,i} \;=\; \mathrm{Trace}\,(Q).$$

Example: $Q = ee^T$, $\mathrm{Trace}\,(Q) = \lambda_{\max}(Q) = n$. In both cases, relative quality is n.

Polyhedral bound

For Boolean $x \in R^n$, we have

$$\langle Qx, x \rangle = \sum_{i,j=1}^{n} Q_{i,j} x_i x_j \leq \sum_{i,j} |Q_{i,j}| \stackrel{\mathrm{def}}{=} \|Q\|_1.$$

How good is it?

Random hyperplane technique. (Krivine 70's, Goemans, Williamson 95)

Let us fix $V \in M_n$. Consider the random vector

$$\xi = \mathrm{sgn}\,[V^T u]$$

with random $u \in R^n$, uniformly distributed on the unit sphere.

([$\cdot$] denotes component-wise operations.)

Lemma 1: $E(\xi_i \xi_j) = \frac{2}{\pi} \arcsin \frac{\langle v_i, v_j \rangle}{\|v_i\| \cdot \|v_j\|}$.

Lemma 2: For $X \succeq 0$, we have $\arcsin[X] \succeq X$.

Proof: $\arcsin[X] = X + \frac{1}{6}[X]^3 + \frac{3}{40}[X]^5 + \ \ldots \ \succeq \ X$.

Quality of polyhedral bound ($Q \succeq 0$)

Let $Q = V^T V$ (this means that $Q_{i,j} = \langle v_i, v_j \rangle$). Then

$$f^*(Q) \geq E(\langle Q\xi, \xi \rangle) \;=\; \frac{2}{\pi} \sum_{i,j=1}^{n} Q^{(i,j)} \arcsin\left(\frac{Q^{(i,j)}}{\sqrt{Q^{(i,i)} Q^{(j,j)}}} \right) \;\stackrel{\text{def}}{=}\; \frac{2}{\pi}\rho.$$

Denote $D = \mathrm{diag}\,(Q)^{-1/2}$. Then $\rho \geq \langle Q, DQD \rangle_M$.

Denote $S_1 = \langle Q, I_n \rangle_M$, $S_2 = \sum\limits_{i \neq j} |Q_{i,j}|$. Then $S_1 + S_2 = \|Q\|_1$.

Thus,

$$\langle Q, DQD \rangle_M = S_1 + \sum_{i \neq j} \frac{(Q_{i,j})^2}{\sqrt{Q_{i,i} Q_{j,j}}} \geq S_1 + \frac{S_2^2}{\sum\limits_{i \neq j} \sqrt{Q_{i,i} Q_{j,j}}}$$

$$= S_1 + \frac{S_2^2}{\left(\sum\limits_{i=1}^{n} \sqrt{Q_{i,i}} \right)^2 - S_1} \geq S_1 + \frac{S_2^2}{nS_1 - S_1} = \|Q\|_1 - S_2 + \frac{S_2^2}{(n-1)(\|Q\|_1 - S_2)}.$$

The minimum is attained for $S_2 = \|Q\|_1 \cdot (1 - \frac{1}{\sqrt{n}})$. Thus,

$$\|Q\|_1 \geq f^*(Q) \geq \langle Q, DQD \rangle_M \geq \frac{2}{1+\sqrt{n}} \|Q\|_1.$$

It is better than the eigenvalue bound!

SDP-bounds: Primal Relaxation (Lovász)

For $X, Y \in M_n$, we have

$$\langle XY, Z\rangle_M = \langle X, ZY^T\rangle_M = \langle Y, X^T Z\rangle_M.$$

Denote $1_n^k : \ (1_n^k)_j = \pm 1, \ j = 1 \ldots n, \ k = 1 \ldots 2^n$.

Then $\langle Q1_n^k, 1_n^k\rangle = \langle Q, 1_n^k(1_n^k)^T\rangle_M$. Therefore

$$f^*(Q) = \max_{X \in \mathcal{P}_n} \langle Q, X\rangle_M,$$

where $\mathcal{P}_n \stackrel{\mathrm{def}}{=} \mathrm{Conv}\,\{1_n^k(1_n^k)^T, \ k = 1 \ldots 2^n\}$. Note that:

- The complete description of $\mathcal{P}_n$ is not known.
- For $X \in \mathcal{P}_n$ we have: $X \succeq 0$, and $d(X) = 1_n$. Thus,

$$f^*(Q) \leq \max\{\langle Q, X\rangle_M : X \succeq 0, d(X) = 1_n\}.$$

Dual Relaxation (Shor)

Problem: $f^*(Q) = \max_x \{\langle Qx, x\rangle : \; x_i^2 = 1, \; i = 1 \dots n\}$.

Its Lagrangian is $\mathcal{L}(x,\xi) = \langle Qx, x\rangle + \sum_{i=1}^{n} \xi_i (1 - (x_i)^2)$. Therefore

$$f^*(Q) = \max_x \min_\xi \mathcal{L}(x,\xi) \leq \min_\xi \max_x \mathcal{L}(x,\xi)$$

$$= \min_\xi \{\langle 1_n, \xi\rangle : \; Q \preceq D(\xi)\} \stackrel{\text{def}}{=} s^*(Q).$$

Note: Both relaxations give exactly the same upper bound:

$$s^*(Q) = \min_\xi \max_{X \succeq 0} \{\langle 1_n, \xi\rangle + \langle X, Q - D(\xi)\rangle_M\}.$$

$$= \max_{X \succeq 0} \min_\xi \{\langle 1_n - D(X), \xi\rangle + \langle X, Q\rangle_M\}.$$

$$= \max_{X \succeq 0} \{\langle X, Q\rangle_M : \; d(X) = 1_n\}.$$

Any hope? (Looks as an attempt to approximate Q by $D(\xi)$.)

Trigonometric form of Quadratic Boolean Problem

We have seen that $f^*(Q) \geq \frac{2}{\pi} \arcsin[V^T V]$ with $d(V^T V) = 1_n$. Let us show that

$$f^*(Q) = \max_{\|v_i\|=1} \frac{2}{\pi} \langle Q, \arcsin[V^T V] \rangle_M .$$

Proof: Choose arbitrary a, $\|a\| = 1$. Let x^* be the global solution.

$$\text{Define } v_i = a \text{ if } x_i^* = 1, \text{ and } v_i = -a \text{ otherwise.}$$

Then $V^T V = x^*(x^*)^T$ and $\frac{2}{\pi} \arcsin[V^T V] = x^*(x^*)^T$. □

Since $\{X = V^T V : d(X) = 1_n\} \equiv \{X \succeq 0 : \; d(X) = 1_n\}$, we get

$$f^*(Q) = \max_{X \succeq 0} \left\{ \frac{2}{\pi} \langle Q, \arcsin[X] \rangle_M : \; d(X) = 1_n \right\}.$$

Corollary: $s^*(Q) \geq f^*(Q) \geq \frac{2}{\pi} s^*(Q)$.

Relative accuracy does not depend on dimension!

General constraints on squared variables

Consider two problems:

$$\phi^* = \max\{\langle Qx, x\rangle : \; [x]^2 \in \mathcal{F}\}, \;\; \phi_* = \min\{\langle Qx, x\rangle : \; [x]^2 \in \mathcal{F}\},$$

where $\mathcal{F}$ is a bounded closed convex set.

Trigonometric form:

$$\phi^* = \max\{\tfrac{2}{\pi}\langle D(d)QD(d), \arcsin[X]\rangle : \\ X \succeq 0, \; d(X) = 1_n, \; d \geq 0, \; [d]^2 \in \mathcal{F}\},$$

$$\phi_* = \min\{\tfrac{2}{\pi}\langle D(d)QD(d), \arcsin[X]\rangle : \\ X \succeq 0, \; d(X) = 1_n, \; d \geq 0, \; [d]^2 \in \mathcal{F}\}.$$

Relaxations:

Define the support function $\xi(u) = \max\{\langle u, v\rangle : \; v \in \mathcal{F}\}$, and

$$\psi^* = \min\{\xi(u) : \; D(u) \succeq Q\}, \;\; \psi_* = \max\{-\xi(u) : \; Q + D(u) \succeq 0\},$$

$$\tau^* = \xi(d(Q)), \;\; \tau_* = -\xi(-d(Q)).$$

Simple relations: $\psi_* \leq \phi_* \leq \tau_* \leq \tau^* \leq \phi^* \leq \psi^*$.

Main result

Denote $\psi(\alpha) = \alpha\psi^* + (1-\alpha)\psi_*$, and $\beta^* = \frac{\psi^*-\tau^*}{\psi^*-\psi_*}$, $\beta_* = \frac{\tau_*-\psi_*}{\psi^*-\psi_*}$.

Theorem. 1. Let

$$\alpha^* = \max\{\tfrac{2}{\pi}\omega(\beta_*), 1-\beta^*\}, \text{ and } \alpha_* = \min\{1-\tfrac{2}{\pi}\omega(\beta^*), \beta_*\},$$

where $\omega(\alpha) = \alpha \arcsin(\alpha) + \sqrt{1-\alpha^2} \quad (\geq 1 + \frac{1}{2}\alpha^2)$.

Then $\quad \psi_* \leq \phi_* \leq \psi(\alpha_*) \leq \psi(\alpha^*) \leq \phi^* \leq \psi^*$.

2. $0 \leq \frac{\phi^*-\psi(\alpha^*)}{\phi^*-\phi_*} \leq \frac{24}{49}$.

3. Define $\bar{\alpha} = \frac{\alpha^*(2-\alpha_*)-\alpha_*}{1+\alpha^*-2\alpha_*}$. Then $\frac{|\phi^*-\psi(\bar{\alpha})|}{\phi^*-\phi_*} \leq \frac{12}{37}$.

Main limitation: Absence of linear constraints

Example. Let $\beta > 0$. Consider the problem

$$\phi^* = \max_x \{\langle Qx, x\rangle : \; [x]^2 = 1_n, \; \langle c, x\rangle = \beta\},$$

$$\phi_* = \min_x \{\langle Qx, x\rangle : \; [x]^2 = 1_n, \; \langle c, x\rangle = \beta\}.$$

Natural relaxation:

$$\psi^* = \max_X \{\langle Q, X\rangle : \; d(X) = 1_n, \; X \succeq 0, \langle Xc, c\rangle = \beta^2\},$$

$$\psi_* = \min_X \{\langle Q, X\rangle : \; d(X) = 1_n, \; X \succeq 0, \langle Xc, c\rangle = \beta^2\}.$$

Denote by v any vector with $[v]^2 = 1_n$.
Assumptions: 1. There exists a unique v_* such that $\langle c, v_*\rangle = \beta$.
2. There exist v_- and v_+ such that $0 < \langle c, v_-\rangle < \beta < \langle c, v_+\rangle$.

Note: in this case $\phi^* = \phi_*$ (unique feasible solution).

Consider the polytope $\mathcal{P}_n = \mathrm{Conv}\,\{V_i = v_i v_i^T,\ i = 1, \ldots, 2^n\}$.

Lemma. Any V_i is an extreme point of $\mathcal{P}_n$. Any pair V_i, V_j is connected by an edge.

Note:

1. In view of our assumption $\exists \tilde{V} \in \mathcal{P}_n$:

$$\tilde{V} = \alpha v_- v_-^T + (1-\alpha) v_+ v_+^T,\ \alpha \in (0,1), \quad \langle \tilde{V} c, c \rangle = \beta^2.$$

2. $\mathcal{P}_n \subset \{X :\ d(X) = 1_n,\ X \succeq 0\}$.

Conclusion: We can choose Q: $\psi^* > \phi^*$.

Since $\psi_* \leq \phi_*$, the relative accuracy of ψ^* is $+\infty$.

Reason of the troubles: We intersect edges of $\mathcal{P}_n$.

This cannot happen if $\beta = 0$.

Further developments

- Boolean quadratic optimization with m homogeneous linear equality constraints (accuracy $O(\ln m)$).
- Quadratic maximization with quadratic inequality constraints (accuracy $O(\ln m)$).

Main bottleneck: absence of cheap relaxations.

Generating functions of integer sets

1. Primal generating functions.

For set $S \subset Z^n$, define $f(S, x) = \sum_{\alpha \in S} x^\alpha$,

where $x^\alpha = \prod_{i=1}^{n} x_i^{\alpha_i}$.

- $f(S, 1_n) = \mathcal{N}(S)$, the *integer volume* of S. Can be used for *counting* problems.
- Sometimes have *short representation*.

Example: $S = \{x \in Z : \ x \geq 0\}$. Then

$$f(S, x) = \frac{1}{1-x}.$$

2. Dual generating functions

2.1. Characteristic function of the set $X \subset Z^n$ is defined as

$$\psi_X(c) = \sum_{x \in X} e^{\langle c,x \rangle}, \text{ if } X \neq \emptyset, \quad \text{and 0 otherwise.}$$

- For counting problem, we have $\mathcal{N}(X) = \psi_X(0)$.
- We can be approximate the optimal value of an optimization problem over X:

$$\mu \ln \psi_X\left(\frac{1}{\mu} c\right) \geq \max_x \{\langle c, x \rangle : \; x \in X(y)\}$$
$$\geq \mu \ln \psi_X\left(\frac{1}{\mu} c\right) - \mu \ln \mathcal{N}(X), \; \mu > 0.$$

2.2. Generating function of *family* $\mathcal{X} = \{X(y), \; y \in \Delta\} \subset Z^m$ is

defined as $\quad g_{\mathcal{X},c}(v) = \sum_{y \in \Delta} \psi_{X(y)}(c) \cdot v^y.$

Dual *counting function:* $f_{\mathcal{X}}(v) = g_{\mathcal{X},0}(v)$.

Hope: short representation. **NB:** Constructed by set *parameters*.

Example

Let $a \in Z^n_+$. Consider the Boolean knapsack polytope

$$B^{1_n}_a(b) = \{x \in \{0,1\}^n : \; \langle a, x \rangle = b\}.$$

Goal: Compute $\mathcal{N}(B^{1_n}_a(b))$ for a given $b \in Z_+$. (It is NP-hard.)

Consider the function $\quad f(z) = \prod\limits_{i=1}^{n} \left(1 + z^{a^{(i)}}\right), \quad$ where

$z \in \mathcal{C} \stackrel{\mathrm{def}}{=} \{z \in C : \; |z| = 1\}$.

We will see later, that $\quad f(z) \equiv \sum\limits_{b=0}^{\|a\|_1} \mathcal{N}(B^{1_n}_a(b))\, z^b, \; z \in \mathcal{C}$,

where $\|a\|_1 \stackrel{\mathrm{def}}{=} \sum\limits_{i=1}^{n} |a^{(i)}|$.

Thus, we need to compute the coefficient of z^b in polynomial $f(z)$.
For that, we compute all previous coefficients.
Direct computation: $\quad O(n \, \|a\|_1) \quad \Rightarrow \quad O(\|a\|_1 \cdot \ln \|a\|_1 \cdot \ln n)$.

Knapsack volumes

Notation: $B_a^u(b) = \{x \in Z^n : \ 0 \le x \le u, \ \langle a, x \rangle = b\}$.

Consider the family $\mathcal{B}_a^u = \{B_a^u(b)\}_{b \in Z_+}$. Its counting function is

$$f_{\mathcal{B}_a^u}(z) \stackrel{\text{def}}{=} \sum_{b=0}^{\infty} \mathcal{N}(B_a^u(b)) \cdot z^b, \qquad z \in \mathcal{C}.$$

Since u is finite, this is a polynomial of degree $\langle a, u \rangle$.

Lemma. $f_{\mathcal{B}_a^u}(z) = \prod_{i=1}^{n} \left(\sum_{k=0}^{u^{(i)}} z^{ka^{(i)}} \right)$.

Proof. For $n = 1$ it is evident.

Denote $a_+ = (a, a^{(n+1)})^T \in Z_+^{n+1}$, and $u_+ = (u, u^{(n+1)})^T \in Z_+^{n+1}$.

For any $b \in Z_+$ we have

$$\mathcal{N}(B_{a_+}^{u_+}(b)) = \sum_{k=0}^{u^{(n+1)}} \mathcal{N}(B_u^a(b - k \cdot a^{(n+1)})).$$

Hence, in view of the inductive assumption, we have

$$\begin{aligned} f_{\mathcal{B}_{a_+}^{u_+}}(z) &= \sum_{b=0}^{\infty} \mathcal{N}(B_{a_+}^{u_+}(b)) \cdot z^b \\ &= \sum_{b=0}^{\infty} \left(\sum_{k=0}^{u^{(n+1)}} \mathcal{N}(B_u^a(b - ka^{(n+1)})) \right) \cdot z^b \\ &= \sum_{b=0}^{\infty} \mathcal{N}(B_u^a(b)) \sum_{k=0}^{u^{(n+1)}} z^{b+ka^{(n+1)}} \\ &= f_{\mathcal{B}_a^u}(z) \cdot \left(\sum_{k=0}^{u^{(n+1)}} z^{ka^{(n+1)}} \right). \quad \square \end{aligned}$$

Complexity

Lemma. Let polynomial $f(z)$ be represented as a product of several polynomials: $\quad f(z) = \prod_{i=1}^{n} p_i(z), \quad z \in \mathcal{C}.$

Then its coefficients can be computed by FFT in

$$O(D(f)\ \ln D(f)\ \ln n\ )$$

arithmetic operations, where $D(f) = \sum_{i=1}^{n} D(p_i)$.

Corollary. All $\langle a, u\rangle$ coefficients of the polynomial $f_{\mathcal{B}_a^u}(z)$ can be computed by FFT in

$$O(\langle a, u\rangle\ \ln\langle a, u\rangle\ \ln n) \text{ a.o.}$$

Unbounded knapsack

Consider $\quad f_{\mathcal{B}_a^\infty}(z) = \sum\limits_{b=0}^{\infty} \mathcal{N}(B_a^\infty(b)) \cdot z^b \; \equiv \; \prod\limits_{i=1}^{n} \frac{1}{1-z^{a^{(i)}}}$,
where $z \in \mathcal{C} \setminus \{1\}$.

Note:

1. The coefficients of the polynomial $g(z) = \prod\limits_{i=1}^{n}(1 - z^{a^{(i)}})$ can be computed by FFT in $O(\|a\|_1 \; \ln \|a\|_1 \; \ln n)$ a.o.

2. After that, the first $b+1$ coefficients of the generating function $f_{\mathcal{B}_a^\infty}(z)$ can be computed in $O(b \; \min\{\ln^2 b, \ln^2 n\})$ a. o.

Generating functions of knapsack polytopes

For *characteristic function* $\psi_X(c) = \sum\limits_{y \in X} e^{\langle c,y \rangle}$ of set X, define its *potential function*: $\quad \phi_X(c) = \ln \psi_X(c)$.

Note that $\quad \xi_X(c) \stackrel{\text{def}}{=} \max\limits_{y \in X} \langle c, y \rangle \leq \phi_X(c) \leq \xi_X(c) + \ln \mathcal{N}(X)$.

Hence, $\quad \xi_X(c) \leq \mu \phi_X(c/\mu) \leq \xi_X(c) + \mu \ln \mathcal{N}(X), \quad \mu > 0$.

For a family of bounded knapsack polytopes $\mathcal{B}_a^u = \{B_a^u(b)\}_{b \in Z_+}$, the generating function looks as follows:

$$g_{\mathcal{B}_a^u, c}(z) = \sum_{b=0}^{\infty} \psi_{B_a^u(b)}(c) \cdot z^b \equiv \sum_{b=0}^{\infty} \exp(\phi_{B_a^u(b)}(c)) \cdot z^b, \quad z \in \mathcal{C}.$$

Short representation: $g_{\mathcal{B}_a^u, c}(z) = \prod\limits_{i=1}^{n} \left(\sum\limits_{k=0}^{u^{(i)}} e^{kc^{(i)}} z^{ka^{(i)}} \right)$.

Unbounded case: $g_{\mathcal{B}_a^\infty, c}(z) = \left[\prod\limits_{i=1}^{n} (1 - e^{c^{(i)}} z^{a^{(i)}}) \right]^{-1}$.

Solving integer knapsack

Find $\quad f^* = \max\limits_{x \in Z^n_+} \{ \langle c, x \rangle : \ \langle a, x \rangle = b \} \ = \ \xi_{B^\infty_a(b)}(c).$

Since f^* is an integer value, we need accuracy less than one.

Note that $\mathcal{N}(B^\infty_a(b)) \leq \prod\limits_{i=1}^{n} \left(1 + \frac{b}{a^{(i)}} \right) \leq (1+b)^n$.

Thus, if we take $\mu < \frac{1}{n} \ln(1+b)$, then

$$-1 + \mu \phi_{B^\infty_a(b)}(c/\mu) < f^* \leq \mu \phi_{B^\infty_a(b)}(c/\mu).$$

For finding coefficient $\psi_{B^\infty_a(b)}(c/\mu) = \exp\{\phi_{B^\infty_a(b)}(c/\mu)\}$, we need

- Compute coefficients of $f(z) = \prod\limits_{i=1}^{n} (1 - e^{c^{(i)}/\mu} \cdot z^{a^{(i)}})$.
- Compute first $b+1$ coefficients of the function $g(z) = \frac{1}{f(z)}$.

This can be done in $O(\|a\|_1 \cdot \ln \|a\|_1 \cdot \ln n \ + \ b \cdot \ln^2 n)$ operations of *exact real arithmetics*.

Further extensions

Problem: count the number of integer points in the set

$$X = \{x \in Z^n : \ 0 \leq x \leq \beta \cdot 1_n, \ Ax = b \in R^m\},$$

where $|A_{i,j}| \leq \alpha$.

Dual counting: $O\left(mn \cdot (1 + \alpha\beta \cdot n)^m\right)$ a.o.

Full enumeration: $O\left(mn \cdot (1 + \beta)^n\right)$ a.o.

For fixed m, the first bound is polynomial in n.