Graver Bases and (Non)-Linear Integer Programming in Polynomial Time

Shmuel Onn

Technion – Israel Institute of Technology

http://ie.technion.ac.il/~onn

Based on several papers joint with several co-authors including
Berstein, De Loera, Hemmecke, Lee, Romanchuk, Rothblum, Weismantel

(Non)-Linear Integer Programming

The problem is:

$$\min/\max \{ f(x) \;:\; Ax = b,\; l \le x \le u,\; x \text{ in } Z^n \}$$

with data:

A: integer m x n matrix

b: right-hand side in Z^m

l,u: lower and upper bounds in Z^n

f: function from Z^n to R

It has generic modeling power and numerous applications

(Non)-Linear Integer Programming

The problem is:

$$\min/\max \{ f(x) \;:\; Ax = b,\;\; l \le x \le u,\;\; x \text{ in } Z^n \}$$

with data:

A: integer $m \times n$ matrix

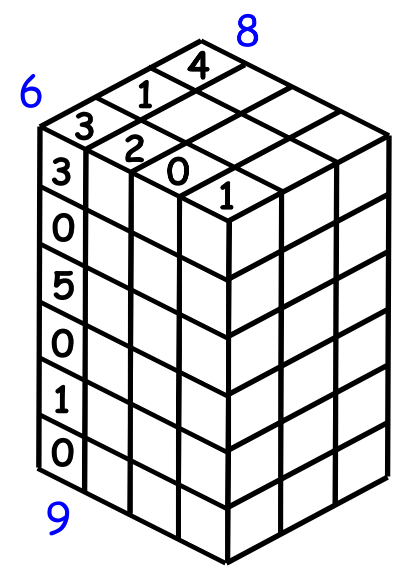
b: right-hand side in Z^m

l,u: lower and upper bounds in Z^n

f: function from Z^n to R

Generic Example: Multiway Tables

Consider (Non)-linear minimization over $l \times m \times n$ tables with given line sums:



It is the integer programming problem:

$$\min \{ f(x) : \Sigma_i x_{i,j,k} = a_{j,k} ,\; \Sigma_j x_{i,j,k} = b_{i,k} ,\; \Sigma_k x_{i,j,k} = c_{i,j} ,\; x \ge 0,\;\; x \text{ in } Z^{l \times m \times n} \}$$

Shmuel Onn

(Non)-Linear Integer Programming

The problem is:

$$\min/\max \{ f(x) \;:\; Ax = b,\; l \le x \le u,\; x \text{ in } Z^n \}$$

with data:

- A: integer $m \times n$ matrix
- b: right-hand side in Z^m
- l,u: lower and upper bounds in Z^n
- f: function from Z^n to R

It has generic modeling power and numerous applications

Unfortunately, even with $f(x)=wx$ linear, it is typically NP-hard

In fixed dimension it is polytime solvable, but often quite limited

We develop new theory enabling polytime solution of broad, natural, universal (non)-linear integer programs in variable dimension

Shmuel Onn

Graver Bases

and

Nonlinear Integer Programming

Graver Bases

The Graver basis of an integer matrix A is the finite set $G(A)$ of conformal-minimal nonzero integer vectors x satisfying $Ax = 0$.

(x is conformal to y if in same orthant and $|x_i| \leq |y_i|$ for all i)

Example: Consider $A=(1\ 2\ 1)$. Then $G(A)$ consists of

circuits: $\pm(2\ {-1}\ 0)$, $\pm(1\ 0\ {-1})$, $\pm(0\ 1\ {-2})$ non-circuits: $\pm(1\ {-1}\ 1)$

Connection to Grobner bases: the set of binomials corresponding to $G(A)$,

$$UGB(A) := \{ x^{v^+} - x^{v^-} \ : \ v \text{ in } G(A) \}$$

forms a universal Grobner basis for the binomial (toric) ideal of A.

Example: for $A=(1\ 2\ 1)$ it is $UGB(A) = \{x_1^2 - x_2,\ x_1 - x_3,\ x_2 - x_3^2,\ x_1 x_3 - x_2\}$

Shmuel Onn

Six Theorems on
(Non)-Linear Integer Programming

Theorem 1: **linear optimization** in **polytime** with **G(A)**:

$$\max \{ wx \ : \ Ax = b, \ \ l \leq x \leq u, \ \ x \text{ in } Z^n \}$$

Reference: N-fold integer programming (De Loera, Hemmecke, Onn, Weismantel)
Discrete Optimization (Volume in memory of George Dantzig), 2008

Shmuel Onn

Six Theorems on
(Non)-Linear Integer Programming

Theorem 2: weighted convex maximization in polytime with $G(A)$:

$$\max \{ f(Wx) : \ Ax = b, \ \ l \le x \le u, \ \ x \text{ in } Z^n \}$$

where W is $d \times n$ matrix and f convex function on Z^d
(balancing d linear criteria or player utilities $W_i x$)

Reference: Convex integer maximization via Graver bases (De Loera, Hemmecke, Onn, Rothblum, Weismantel) Journal of Pure and Applied Algebra, 2009

Shmuel Onn

Six Theorems on
(Non)-Linear Integer Programming

Theorem 3: **separable convex minimization** in **polytime** with **G**(**A**):

$$\min \{ \sum f_i(x_i) \; : \; Ax = b, \;\; l \le x \le u, \;\; x \text{ in } Z^n \}$$

Reference: A polynomial oracle-time algorithm for convex integer minimization (Hemmecke, Onn, Weismantel) Mathematical Programming, to appear

Six Theorems on
(Non)-Linear Integer Programming

Theorem 4: **integer point l_p-nearest to** x in **polytime** with $G(A)$:

$$\min \{ |x - x|_p \; : \; Ax = b, \quad l \leq x \leq u, \quad x \text{ in } Z^n \}$$

Reference: A polynomial oracle-time algorithm for convex integer minimization (Hemmecke, Onn, Weismantel) Mathematical Programming, to appear

Six Theorems on
(Non)-Linear Integer Programming

Theorem 5: **quadratic minimization** in polytime with **$G(A)$**:

$$\min \{x^T V x \ : \ Ax = b, \ \ l \leq x \leq u, \ \ x \text{ in } Z^n\}$$

where V lies in cone $K_2(A)$ of possibly indefinite matrices, enabling minimization of some convex and some non-convex quadratics.

Reference: The quadratic Graver cone, quadratic integer minimization & extensions (Lee, Onn, Romanchuk, Weismantel), submitted

Six Theorems on
(Non)-Linear Integer Programming

Theorem 6: **polynomial minimization** in polytime with **G(A)**:

$$\min \{p(x) \; : \; Ax = b, \quad l \le x \le u, \quad x \text{ in } Z^n\}$$

where p is possibly indefinite polynomial of degree d in cone $K_d(A)$,
enabling minimization of some (non)-convex degree d polynomials.

Reference: The quadratic Graver cone, quadratic integer minimization & extensions
(Lee, Onn, Romanchuk, Weismantel), submitted

Some Proofs

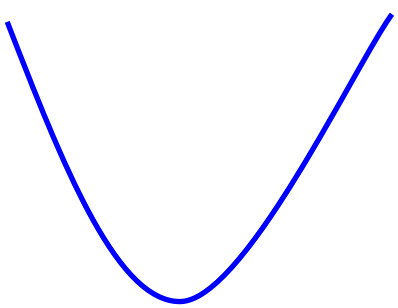
Proof of Theorem 3
(separable convex minimization)

Lemma 1: Any separable convex function f on R^n is supermodular, that is, for any vectors g_i in the same orthant and any vector x, it satisfies

$$f(x+\sum g_i) - f(x) \geq \sum \left(f(x+g_i) - f(x)\right)$$

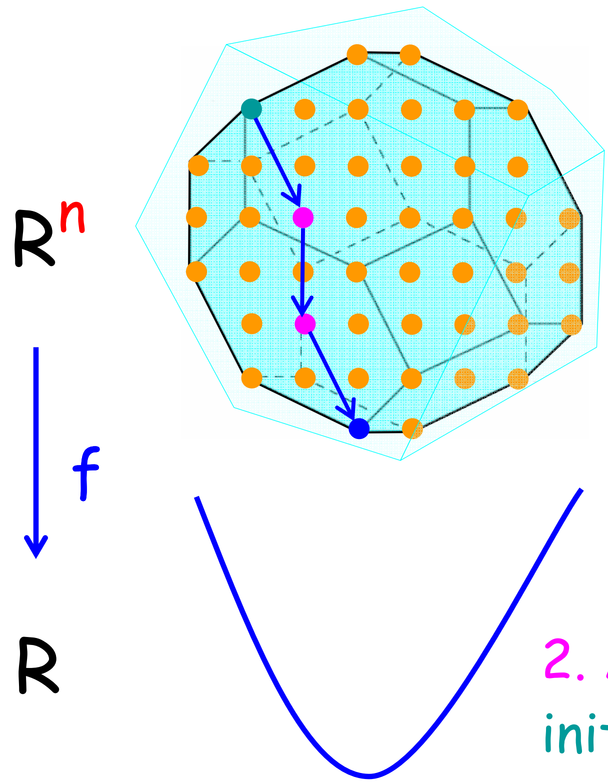
Lemma 2: For separable convex f, point x, bounds l,u and direction g in R^n, the following univariate integer program can be solved in polytime:

$$\min \{ f(x + \alpha g) \;:\; l \leq x + \alpha g \leq u, \quad \alpha \text{ nonnegative integer} \}$$



Shmuel Onn

Proof of Theorem 3
(separable convex minimization)



Solve $\min\{\sum f_i(x_i) : Ax = b,\ l \leq x \leq u,\ x \text{ in } Z^n\}$

using the Graver basis $G(A)$, as follows:

1. Find initial point by auxiliary program

2. Apply Lemma 2 repeatedly to greedily augment initial point to optimal one using directions g in $G(A)$

Using the supermodularity of f from Lemma 1 and integer Caratheodory theorem get polynomial time convergence

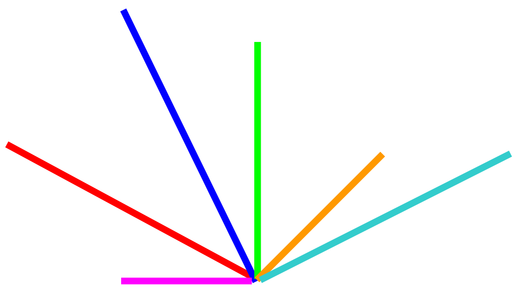
Proof of Theorem 1: linear function $wx=\sum w_i x_i$: special case of Theorem 3

Proof of Theorem 2
(weighted convex maximization)

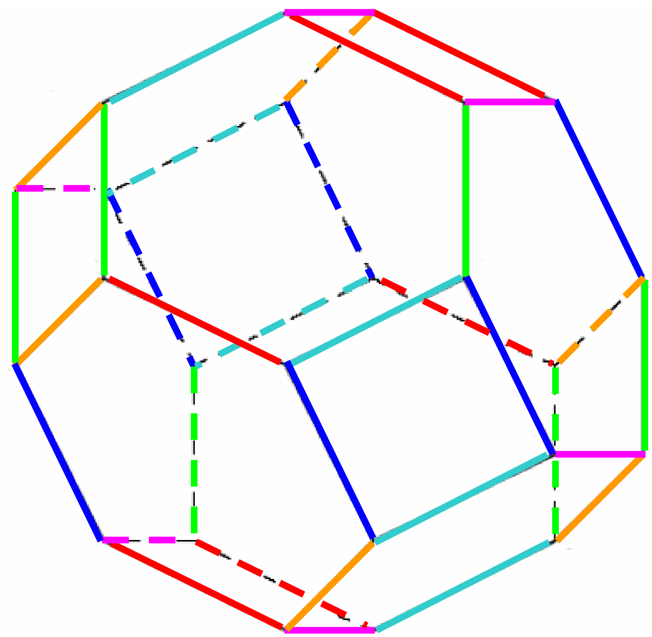
Lemma: Linear optimization over S in Z^n can be used to solve in polytime

$$\max \{ f(Wx) \ : \ x \text{ in } S \}$$

provided we are given a set E of all edge-directions of conv(S)



set E of all edge-directions of conv(S)



Proof of Theorem 2
(weighted convex maximization)

Lemma: Linear optimization over S in Z^n can be used to solve in polytime

$$\max \{ f(Wx) \ : \ x \text{ in } S \}$$

Proof of Theorem 2:

Given $S := \{x \text{ in } Z^n : Ax = b, \ l \leq x \leq u\}$ and the Graver basis $G(A)$, do:

1. Use the Graver basis as set $E:=G(A)$ of all edge-directions of conv(S)

2. Use $G(A)$ for linear-optimization over S via Theorem 1

3. Apply Lemma for weighted convex maximization, repeatedly using 2.

Shmuel Onn

N-Fold Integer Programming

N-Fold Products

The n-fold product of an $(r,s) \times t$ bimatrix A is the following $(r+ns) \times nt$ matrix:

$$A^{(n)} = \underbrace{\begin{pmatrix} A_1 & A_1 & A_1 & \cdots & A_1 \\ A_2 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_2 \end{pmatrix}}_{n} .$$

Graver Bases of N-Fold Products

Lemma: For fixed A, can compute in polytime the Graver basis $G(A^{(n)})$ of

$$A^{(n)} = \begin{pmatrix} A_1 & A_1 & A_1 & \cdots & A_1 \\ A_2 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_2 \end{pmatrix} .$$

The proof uses finiteness results of

Aoki-Takemura, Santos-Sturmfels, Hosten-Sullivant

(Non)-Linear N-Fold Integer Programming

Theorem: we can solve each of the following in **polynomial time**:

linear optimization: $\max\{wx : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

weighted convex maximization: $\max\{f(Wx) : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

separable convex minimization: $\min\{\sum f_i(x_i) : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

$$A^{(n)} = \underbrace{\begin{pmatrix} A_1 & A_1 & A_1 & \cdots & A_1 \\ A_2 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_2 \end{pmatrix}}_{n}$$

References: Theory and applications of n-fold integer programming, 35 pages, IMA Volume on Mixed Integer Nonlinear Programming, Springer, to appear

(Non)-Linear N-Fold Integer Programming

Theorem: we can solve each of the following in **polynomial time**:

linear optimization: $\max\{wx : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

weighted convex maximization: $\max\{f(Wx) : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

separable convex minimization: $\min\{\sum f_i(x_i) : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

Proof: Use Lemma to construct in polytime the Graver base $G(A^{(n)})$.

Now apply and use Theorems 1 - 3 to optimize in polytime.

(Non)-Linear N-Fold Integer Programming

Theorem: we can solve each of the following in **polynomial time**:

linear optimization: $\max\{wx : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

weighted convex maximization: $\max\{f(Wx) : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

separable convex minimization: $\min\{\sum f_i(x_i) : A^{(n)}x = b,\ l \le x \le u,\ x \text{ in } Z^{nt}\}$

With more work can also do weighted separable convex minimization:

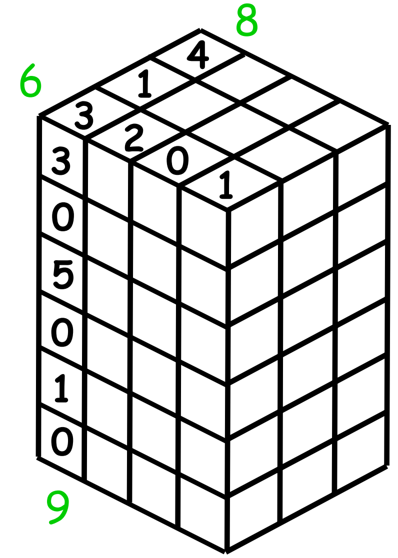
$$\min\{f(W^{(n)}x) : A^{(n)}x = b,\ l \le x \le u,\ L \le W^{(n)}x \le U,\ x \text{ in } Z^{nt}\}$$

Some Applications

1. Multiway Tables

Complexity of deciding the existence of

$l \times m \times n$ tables with given line sums:



- l, m, n variable: NP-complete

  Three dimensional matching, Karp, 1972

- l fixed, m, n variable: Universal for IP (even with l=3)

  Part of my talk in previous Japan GB conference, De Loera, Onn, 2006

- l,m fixed, n variable: Polytime

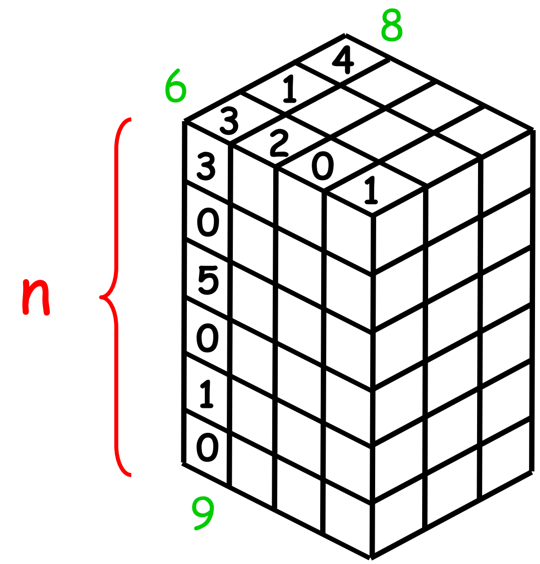
  Consequence of linear n-fold IP, De Loera, Hemmecke, Onn, Weismantel, 2008

- l, m, n fixed: Polytime

  Integer programming in fixed dimension, Lenstra, 1982

1. Multiway Tables

Much more generally, consider the multi-index transportation problem studied by Motzkin in 1952, of minimization over $m_1 \times \cdots \times m_k \times n$ tables with given margins:



It is an n-fold program

$$\min\{f(x) \;:\; A^{(n)}x = b, \quad x \geq 0, \quad x \text{ integer}\}$$

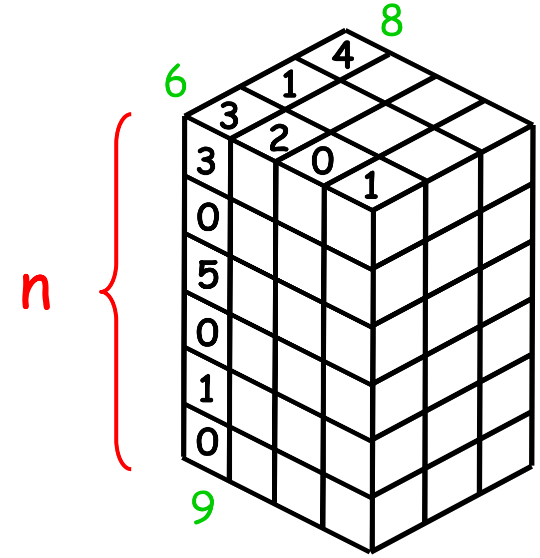
$$A^{(n)} = \underbrace{\begin{pmatrix} A_1 & A_1 & A_1 & \cdots & A_1 \\ A_2 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_2 \end{pmatrix}}_{n}$$

for suitable A depending on $m_1, \ldots, m_k$ where:

- A_1 gives equations of margins summing over layers
- A_2 gives equations of margins summing within a single layer at a time

1. Multiway Tables

Much more generally, consider the multi-index transportation problem studied by Motzkin in 1952, of minimization over $m_1 \times \cdots \times m_k \times n$ tables with given margins:



Corollary 1: (Non)-linear optimization over $m_1 \times \cdots \times m_k \times n$ tables with given margins can be done in **polynomial time**

In contrast: **Universality of three-way tables** (De Loera, Onn):
Every integer program is one over $3 \times m \times n$ tables with given line-sums

2. Privacy in Statistical Data Bases

Common strategy in web disclosure of sensitive data:
disclose margins but not table entries.

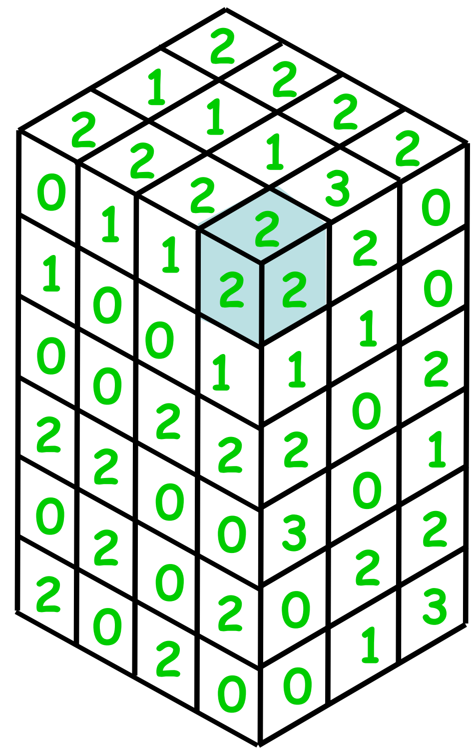
The security of an entry is then related to the set of values
that it can take in all tables with the disclosed margins.

2. Privacy in Statistical Data Bases

Universality of Table Entries (My talk in previous Japan GB Conference):

Every finite set of nonnegative integers is the set of values in an entry of the $3 \times m \times n$ tables with some given line-sums

Example: the values occurring in the shaded entry in the tables with the given line-sums are precisely 0, 2



Shmuel Onn

2. Privacy in Statistical Data Bases

In contrast, the theory of n-fold integer programming yields:

Corollary 2: The set of values in any entry in all $m_1 \times \cdots \times m_k \times n$ tables with any given margins can be computed in polytime

Proof: compute the true integer lower and upper bounds on the entry by solving the following two n-fold programs in polytime:

$L = \min\ x_{i_1 \ldots i_{k+1}}$ over all tables with the given margins

$U = \max\ x_{i_1 \ldots i_{k+1}}$ over all tables with the given margins

(note that the value is unique if and only if $L = U$)

Incorporate bounds $L+1 \leq x_{i_1 \ldots i_{k+1}} \leq U-1$ and repeat.

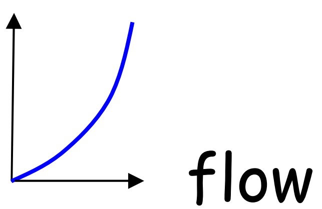
3. Multicommodity Flows

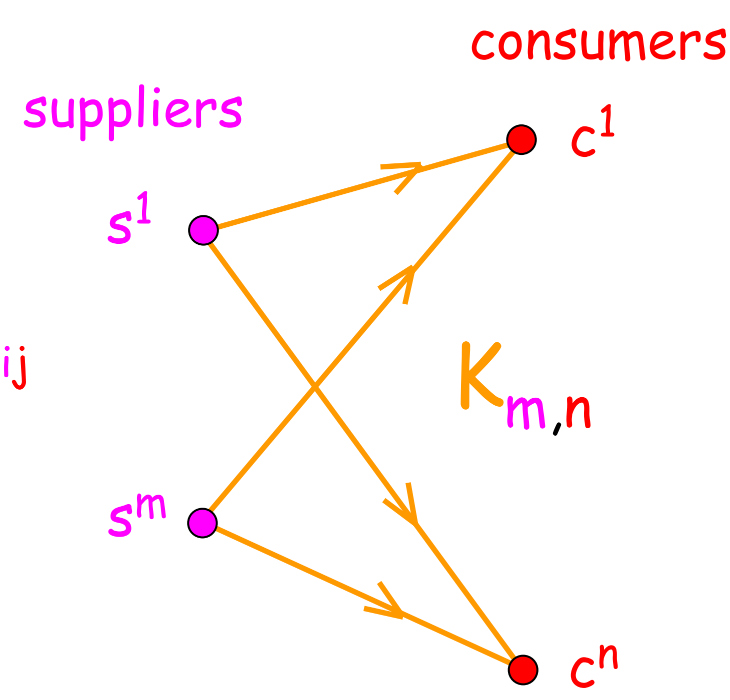
Find integer l-commodity flow x from m suppliers to n consumers under supply, consumption and capacity constraints, of minimum possibly convex cost f which accounts for channel congestion

It can be shown to be a (non)-linear n-fold integer program

$$\min \{ f(W^{(n)}x) \ : \ A^{(n)}x = (s^i, c^j), \ \ x \geq 0, \ \ W^{(n)}x \leq u, \ \ x \text{ in } Z^{mnl} \}$$

cost: $\Sigma_{ij} c_{ij} (\Sigma_k x_{ijk})^{a_{ij}}$





capacities: u_{ij}

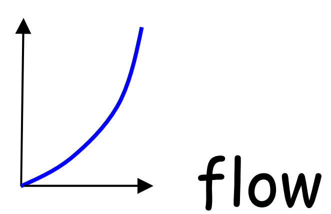
Shmuel Onn

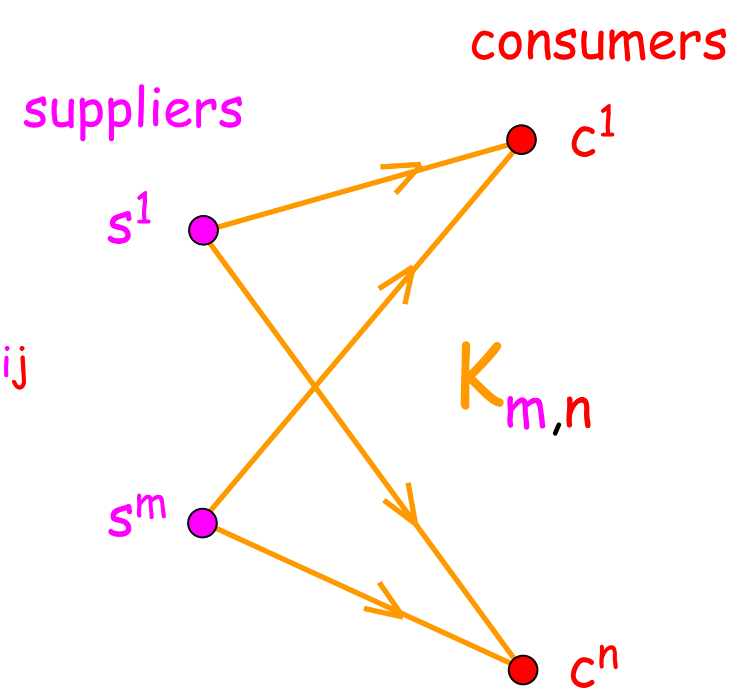
3. Multicommodity Flows

Find integer l-commodity flow x from m suppliers to n consumers under supply, consumption and capacity constraints, of minimum possibly convex cost f which accounts for channel congestion

Corollary 3: For any fixed l commodities and m suppliers, can find optimal multicommodity flow for n consumers in polytime

cost: $\sum_{ij} c_{ij} (\sum_k x_{ijk})^{a_{ij}}$





capacities: u_{ij}

Shmuel Onn

4. Stochastic Integer Programming

In this important model, part of the data is random, and decisions are in two stages - x before and y after the realization of random data:

$$\min \{ wx + E[c(x)] \ : \ x \geq 0, \ x \text{ in } Z^r \}$$

where

$$c(x) = \min \{ uy \ : \ A_1 x + A_2 y = b, \ y \geq 0, \ y \text{ in } Z^s \}$$

Suitably discretizing the sample space into n scenarios, the problem becomes a transposed n-fold integer program.

While the Graver basis here cannot be computed in polytime, with some extra work we do get the following:

Corollary 4: Stochastic IP with n scenarios can be solved in polytime

Universality

Universality of N-Fold Integer Programming

Consider the following special form of the n-fold product operator,

$$A^{[n]} = \underbrace{\begin{pmatrix} I & I & I & \cdots & I \\ A & 0 & 0 & \cdots & 0 \\ 0 & A & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A \end{pmatrix}}_{n} .$$

Consider such m-fold products of the 1×3 matrix $[1\ \ 1\ \ 1]$. For example,

$$[1\ \ 1\ \ 1]^{[3]} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} .$$

Universality of N-Fold Integer Programming

$$A^{[n]} = \begin{pmatrix} I & I & I & \cdots & I \\ A & 0 & 0 & \cdots & 0 \\ 0 & A & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A \end{pmatrix} .$$

Universality Theorem: Any bounded set $\{ y \text{ integer} : By = b, y \geq 0 \}$ is in polynomial-time-computable coordinate-embedding-bijection with some

$$\{ x \text{ integer} : [1\ \ 1\ \ 1]^{[m][n]} x = a,\ x \geq 0 \}$$

Reference: *All linear and integer programs are slim 3-way programs* (De Loera, Onn) SIAM Journal on Optimization

Shmuel Onn

Universality of N-Fold Integer Programming

$$A^{[n]} = \begin{pmatrix} I & I & I & \cdots & I \\ A & 0 & 0 & \cdots & 0 \\ 0 & A & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A \end{pmatrix} .$$

Universality Theorem: Any bounded set $\{ y \text{ integer} : By = b, y \geq 0 \}$ is in polynomial-time-computable coordinate-embedding-bijection with some

$$\{ x \text{ integer} : [1\ \ 1\ \ 1]^{[m][n]} x = a,\ \ x \geq 0 \}$$

Scheme for Nonlinear Integer Programming:

any integer program $\max \{ f(y) : By = b,\ y \geq 0,\ y \text{ integer} \}$

can be lifted to:

n-fold program: $\max \{ f(x) : [1\ 1\ 1]^{[m][n]} x = a,\ x \geq 0,\ x \text{ integer} \}$

Shmuel Onn

Epilogue:

Nonlinear Discrete Optimization

Setup for Nonlinear Discrete Optimization

The problem is:

$$\min/\max \{ f(Wx) : x \text{ in } S \}$$

with S set in Z^n, W integer $d \times n$ matrix, f function on Z^d.

It can be interpreted as balancing d criteria or player utilities $W_i x$ and enables determination of broad useful classes of triples S,W,f solvable efficiently (deterministically, randomly, or approximately)

Setup for Nonlinear Discrete Optimization

The problem is:

$$\min/\max \{ f(Wx) : x \text{ in } S \}$$

with S set in Z^n, W integer $d \times n$ matrix, f function on Z^d.

The presentation of S induces two branches:

Integer Programming:

$S = \{x \text{ in } Z^n : A(x) \leq 0 \}$

given by (non)-linear inequalities

Combinatorial Optimization:

S in $\{0,1\}^n$

given compactly or by oracle

Three Nonlinear Combinatorial Optimization Examples

The problem is:

$$\min/\max \{ f(Wx) : x \text{ in } S \}$$

Theorem A: For S **matroid** (e.g. trees, experimental designs) in **polytime**.

Berstein, Lee, Maruri-Aguilar, Onn, Riccomagno, Weismantel, Wynn, SIAM J. Disc. Math.

Yael

Hugo

Three Nonlinear Combinatorial Optimization Examples

The problem is:

$$\min/\max \{ f(Wx) : x \text{ in } S \}$$

Theorem A: For S **matroid** (e.g. trees, experimental designs) in **polytime**.

Berstein, Lee, Maruri-Aguilar, Onn, Riccomagno, Weismantel, Wynn, SIAM J. Disc. Math.

Theorem B: For S **matroid intersection** in **randomized polytime**.

Berstein, Lee, Onn, Weismantel, Mathematical Programming, to appear

Theorem C: For S **independence system**, d=1, **approximation in polytime**.

Lee, Onn, Weismantel, SIAM J. Disc. Math.

Bibliography (mostly available at http://ie.technion.ac.il/~onn)

Theory and applications of n-fold integer programming, 35 pages,
IMA Volume on Mixed Integer Nonlinear Programming, Springer, to appear

- Convex matroid optimization (SIAM Disc. Math.)
- The complexity of 3-way tables (SIAM J. Comp.)
- Convex combinatorial optimization (Disc. Comp. Geom.)
- Markov bases of 3-way tables (J. Symb. Comp.)
- All linear and integer programs are slim 3-way programs (SIAM J. Opt.)
- Entry Uniqueness in margined tables (Lect. Notes Comp. Sci.)
- Graver complexity of integer programming (Annals Combin.)
- N-fold integer programming (Disc. Opt. in memory of Dantzig)
- Convex integer maximization via Graver bases (J. Pure App. Algebra)
- Polynomial oracle-time convex integer minimization (Math. Prog.)
- Nonlinear matroid optimization and experimental design (SIAM Disc. Math.)
- Nonlinear optimization over a weighted independence system (SIAM Disc. Math.)
- Nonlinear optimization for matroid intersection and extensions (Math. Prog.)
- N-fold integer programming and nonlinear multi-transshipment (submitted)
- The quadratic Graver cone, quadratic integer minimization & extensions (submitted)

Comprehensive treatment is in my new monograph:

Nonlinear Discrete Optimization:
An Algorithmic Theory

**Zurich Lectures in Advanced Mathematics,
European Mathematical Society, 150 pages, to appear**

**Based on my Nachdiplom Lectures
delivered at ETH Zurich in Spring 2009
(preliminary notes are in my homepage)**