User's Guide for 4ti2 version 1.3.1

A software package for algebraic, geometric and combinatorial problems on linear spaces

January 5, 2007

Contents

1	\mathbf{Beg}	inner'	s guide	5
	1.1	Linear	systems and their encodings	5
		1.1.1	Linear systems and integer linear systems	5
		1.1.2	Specifying a linear system in 4ti2	6
		1.1.3	How does an explicit solution to linear systems look like?	7
	1.2	Brief t	cutorial	8
		1.2.1	Solving linear systems over $\mathbb R$ and over $\mathbb Z$	8
		1.2.2	Computing extreme rays and Hilbert bases	12
		1.2.3	Computing circuits and Graver bases	15
		1.2.4	Integer programming and toric Gröbner bases	18
		1.2.5	Markov Bases in Statistics	20

CONTENTS

Chapter 1

Beginner's guide

In this part, we use a few sample problems to introduce you to the basic functionality of 4ti2. After working through this part, you should know about *linear systems* and their encodings in 4ti2, and should be able to do computations using the following functions:

- qsolve, rays, circuits
- zsolve, hilbert, graver, ppi
- minimize, groebner, normalform
- genmodel, markov

1.1 Linear systems and their encodings

In this section you learn about the data structure linear system and how it is specified in 4ti2.

1.1.1 Linear systems and integer linear systems

In 4ti2, a linear system is defined by d constraints $Ax \sim b$ in n unknowns x, where each constraint is either \leq , = or \geq , that is $\sim \in \{\leq, =, \geq\}^d$. Moreover, one may

specify sign constraints on the variables that need to be respected in an explicit continuous/integer representation of all solutions.

There is no particular difference in 4ti2 between a linear system and an integer linear system. Currently, the user chooses between one of the two by calling the appropriate functions on the linear system.

1.1.2 Specifying a linear system in 4ti2

In order to use a linear system as input, we need to specify its parts to 4ti2. As our running example, take

$$\left(\begin{array}{ccc} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{array}\right) x \stackrel{\leq}{\leq} \left(\begin{array}{c} 6 \\ 10 \end{array}\right)$$

with sign constraints (1, 2, 2, 0), which we will explain below.

First, we have to give our problem a project name, say PROJECT.

- The matrix A has to be put into the file PROJECT.mat.
 - 2 4
 - 1 1 1 1
 - 1 2 3 4
- The relations \sim then have to be specified in PROJECT.rel.
 - 1 2
 - < <
- The right-hand side vector goes into PROJECT.rhs.
 - 1 2
 - 6 10
- And finally, the sign constraints end up in PROJECT.sign.
 - 1 4
 - 1 2 2 0

Note.

- The input files all have the format of a matrix, preceded by the matrix dimensions. As the dimensions already specify how many symbols have to be read, the matrix could also be given in only one line or even in many lines of different lengths.
- In 4ti2 version 1.3.1, all appearing numbers have to be integers.
- Consequently, this implies that, at the moment, qsolve only supports homogeneous linear systems, that is systems with b = 0, since minimal inhomogeneous solutions could have rational components.

1.1.3 How does an explicit solution to linear systems look like?

If the system is solved over \mathbb{R} (using qsolve), 4ti2 returns two sets of integer vectors:

- a set H of support-minimal homogeneous solutions, and
- a set F defining the linear vector space the solution set lives in.

As only homogeneous linear systems are supported in this version of 4ti2, no list of minimal inhomogeneous solutions is computed. Any solution z of the linear system can now be written as

$$z = \sum \alpha_j h_j + \sum \beta_k f_k \tag{1.1}$$

with $h_j \in H$, $f_k \in F$, and $\alpha_j \ge 0$.

If the system is solved over \mathbb{Z} (using zsolve), 4ti2 returns three sets of integer vectors:

- a set *H* of minimal homogeneous *integer* solutions,
- a set I of minimal inhomogeneous integer solutions, and
- a set F defining the sublattice of \mathbb{Z}^n the solution set lives in.

Any solution z of the linear system can now be written as

$$z = i + \sum \alpha_j h_j + \sum \beta_k f_k \tag{1.2}$$

for some $i \in I$ and with $h_j \in H$, $f_j \in F$, and $\alpha_j \in \mathbb{Z}_+$.

Sign file. Let us finally clarify what the sign file PROJECT.sign is good for. The sign file may declare a variable to be non-negative (1), to be non-positive (-1), or to consider both cases independently and unite the answers (2). If a nonzero sign has been assigned to a variable, the explicit representations (1.1) and (1.2) above of a solution z have to respect the sign on that variable. The default setting for each variable is 0 (when using qsolve and zsolve), that is, the sign need not be respected in the explicit representation. In our example above, the first variable is declared to be non-negative, the second and the third one expand to $2 \cdot 2 = 4$ orthant constraints, and the fourth variable is unconstrained. Note, however, that 4 ti 2 does not decompose the problem internally into the four problems with sign patterns (1, 1, 1, 0), (1, 1, -1, 0), (1, -1, 1, 0), and (1, -1, -1, 0), but deals with them more efficiently at the same time.

1.2 Brief tutorial

1.2.1 Solving linear systems over $\mathbb R$ and over $\mathbb Z$

In this example you learn about the functions qsolve and zsolve.

Let us have a look at the linear system

$$\begin{array}{rcl}
x & - & y & \leq & 2 \\
-3x & + & y & \leq & 1 \\
x & + & y & \geq & 1 \\
y & \geq & 0
\end{array}$$

and let us solve it over \mathbb{R} , we have to create the files encoding the linear system. Let us call our project system. Then the input files look as follows:

system.mat	system.rel	system.rhs	system.sign		
3 2	1 3	1 3	1 2		
1 -1	< < >	2 1 1	0 1		
-3 1					
1 1					

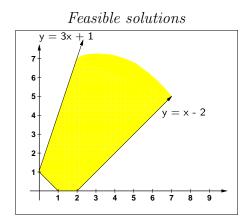
and then call

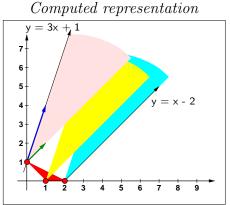
./qsolve system

This call creates three files

system.qinhom	system.qhom	system.qfree
3 2	2 2	0 2
0 1	1 1	
0 2	1 3	
1 0		

which correspond to the explicit description of all solutions:





$$\operatorname{conv}\left(\binom{1}{0}, \binom{2}{0}, \binom{0}{1}\right) + \operatorname{cone}\left(\binom{1}{1}, \binom{1}{3}\right)$$

Note that in the second picture above, the three colored cones are only a simplification and shall visualize the covering of the feasible region by infinitely many shifted copies of the cone

$$\operatorname{cone}\left(\binom{1}{1},\binom{1}{3}\right),$$

one appended to each point in

$$\operatorname{conv}\left(\binom{1}{0}, \binom{2}{0}, \binom{0}{1}\right).$$

Let us now turn to the integer situation, that is, let us solve the system

$$\begin{array}{rcl}
x & - & y & \leq & 2 \\
-3x & + & y & \leq & 1 \\
x & + & y & \geq & 1 \\
y & \geq & 0
\end{array}$$

over \mathbb{Z} . As the linear system itself is unchanged, we can use the same input files as above in order to specify the linear system.

system.mat	system.rel	system.rhs	system.sign		
3 2	1 3	1 3	1 2		
1 -1	< < =	2 1 1	0 1		
-3 1					
1 1					

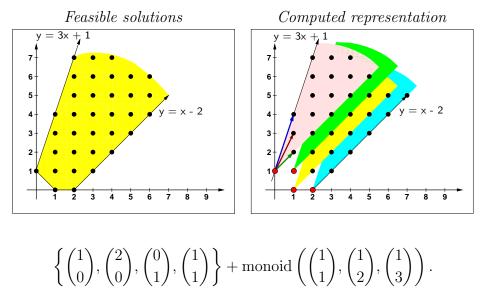
Then, however, we call

./zsolve system

This call creates three files

system.zinhom	system.zhom	system.zfree
4 2	3 2	0 2
0 1	1 1	
0 2	1 2	
1 0	1 3	
1 1		

which correspond to the explicit description of all integer solutions:



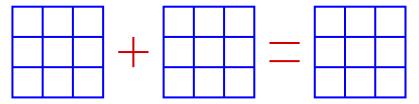
Note that in the pictures above, we are only interested in the *lattice points* inside the colored regions! The full regions are colored only for the purpose of visualizing the covering of all feasible integer solutions by finitely many shifted copies of the monoid

monoid
$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \end{pmatrix}$$
.

1.2.2 Computing extreme rays and Hilbert bases

In this example you learn about the functions rays and hilbert.

Let us consider the set of magic 3×3 squares with non-negative real entries, that is, the set of all 3×3 arrays with non-negative real entries whose row sums, column sums, and main diagonal sums all add up to the same number, the magic constant of the square.



Clearly, addition of two magic squares gives another magic square, as well as does multiplication of a magic square by a non-negative number. Therefore, we may talk about the *cone* of magic 3×3 squares. In fact, this cone is a pointed rational

polyhedral cone described by the linear system

$$x_{11} + x_{12} + x_{13} = x_{21} + x_{22} + x_{23}$$

$$= x_{31} + x_{32} + x_{33}$$

$$= x_{11} + x_{21} + x_{31}$$

$$= x_{12} + x_{22} + x_{32}$$

$$= x_{13} + x_{23} + x_{33}$$

$$= x_{11} + x_{22} + x_{33}$$

$$= x_{31} + x_{22} + x_{13}$$

$$x_{ij} \ge 0, \text{ for all } i, j = 1, 2, 3.$$

Bringing all x_{ij} to the left-hand side of these equations, the matrix $A_{3\times 3}$ defining this linear system is

$$A_{3\times3} = \begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & -1 & 0 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 & -1 & 0 & -1 & 0 & 0 \end{pmatrix}.$$

Below, we will deal with the more interesting case of *integer* magic squares. For the moment, however, we wish to compute the extreme rays of the magic square cone $\{z: A_{3\times 3}z = 0, z \geq 0\}.$

In order to call the function rays, we only have to create one file, say magic3x3.mat, in which we specify the problem matrix $A_{3\times3}$. The remaining data is set by default to "equations only", to "homogeneous system", and to "all variables are non-negative". Note that we are allowed to change these defaults (except homogeneity) by specifying data in magic3x3.rel and magic3x3.sign

mag	magic3x3.mat								
7	9								
1	1	1	-1	-1	-1	0	0	0	
1	1	1	0	0	0	-1	-1	-1	
0	1	1	-1	0	0	-1	0	0	
1	0	1	0	-1	0	0	-1	0	
1	1	0	0	0	-1	0	0	-1	
0	1	1	0	-1	0	0	0	-1	
1	1	0	0	-1	0	-1	0	0	

Now we call

./rays magic3x3

which creates the single file

magic3x3.ray									
4	9								
0	2	1	2	1	0	1	0	2	
1	2	0	0	1	2	2	0	1	
2	0	1	0	1	2	1	2	0	
1	0	2	2	1	0	0	2	1	

that corresponds to the four extremal rays of the 3×3 magic square cone:

0	2	1	1	2	0	2	0	1	1	0	2
2	1	0	0	1	2	0	1	2	2	1	0
1	0	2	2	0	1	1	2	0	0	2	1

Every magic 3×3 square is a non-negative linear combination of these four magic squares.

If we turn now to *integer* magic squares, we are looking for a Hilbert basis of the 3×3 magic square cone. As the default settings for hilbert are the same as for rays, we can use the same input file

mag	magic3x3.mat								
7	9								
1	1	1	-1	-1	-1	0	0	0	
1	1	1	0	0	0	-1	-1	-1	
0	1	1	-1	0	0	-1	0	0	
1	0	1	0	-1	0	0	-1	0	
1	1	0	0	0	-1	0	0	-1	
0	1	1	0	-1	0	0	0	-1	
1	1	0	0	-1	0	-1	0	0	

for this computation. However, to compute the Hilbert basis, we call

./hilbert magic3x3

which creates the single output file

magic3x3.hil									
4	9								
0	2	1	2	1	0	1	0	2	
1	2	0	0	1	2	2	0	1	
2	0	1	0	1	2	1	2	0	
1	0	2	2	1	0	0	2	1	
1	1	1	1	1	1	1	1	1	

that corresponds to the five elements in the minimal Hilbert basis of the 3×3 magic square cone:

0	2	1	
2	1	0	
1	0	2	

1	2	0
0	1	2
2	0	1

2	0	1
0	1	2
1	2	0

1	0	2
2	1	0
0	2	1

1	1	1
1	1	1
1	1	1

Every integer magic 3×3 square is a non-negative *integer* linear combination of these five integer magic squares. Note that the all-1 square is in the interior of the magic square cone.

1.2.3 Computing circuits and Graver bases

In this example you learn about the functions graver, ppi, and circuits.

As an example of a Graver basis computation, let us compute the primitive partition identities of order n = 4. Before we do the simple computation, let us explain what a primitive partition identity is.

A partition identity is any identity of the form

$$a_1 + \ldots + a_k = b_1 + \ldots + b_l$$

with (generally not distinct) integer numbers $0 < a_i, b_j \le n$. A partition identity is called **primitive** if no proper subidentity exists.

For example,

$$1+2+3=2+2+2$$

is a partition identity which is not primitive, since it contains the subidentity

$$1+3=2+2$$

which is in fact primitive.

The description of the primitive partition identities for fixed n, however, is exactly the description of the Graver basis of the matrix

$$A_n = \left(\begin{array}{ccccc} 1 & 2 & 3 & \dots & n \end{array} \right).$$

Let us finally do the computation for n = 3. We create an input file ppi3 for 4ti2 which looks as follows:

ppi3	.mat	,
1	3	
1	2	3

and call

./graver ppi3

This call will create an output file ppi3.gra that looks like:

ppi3	ppi3.gra						
5	3						
3	0	-1					
2	-1	0					
0	3	-2					
1	1	-1					
1	-2	1					

Thus, there are 5 primitive partition identities of order n=3:

$$1+1+1 = 3$$

$$1+1 = 2$$

$$2+2+2 = 3+3$$

$$1+2 = 3$$

$$1+3 = 2+2$$

You may try and compute the primitive partition identities for bigger n, say n = 17, 20, or 23. Be aware, especially the latter two problems take a long, long time. What is the biggest n for which you can compute the primitive partition identities of order n on your machine within one hour?

Due to the very special structure of the matrix, there are algorithmic speed-ups [1, 2, 4]. The currently fastest algorithm to compute primitive partition identities is implemented in the function ppi of 4ti2. Try running

which creates two files ppi17.mat (so we do not really have to create this file ourselves) and the file ppi17.gra containing the desired identities. Compare this running time with the time taken by

./graver ppi17

Do you notice the speed-up?

Let us now turn to the question of determining the support-minimal partition identities. This, in fact, is the question of computing the circuits of the matrix

$$A_n = \left(\begin{array}{ccccc} 1 & 2 & 3 & \dots & n \end{array} \right).$$

We use the same input file

ppi3.mat					
1	3				
1	2 3				

as above and call

./circuits ppi3

This call will create an output file ppi3.cir that looks like:

ppi3.cir						
3	3					
3	0	-1				
2	-1	0				
0	3	-2				

Thus, there are 3 support-minimal partition identities of order n = 3:

$$1+1+1 = 3$$

 $1+1 = 2$
 $2+2+2 = 3+3$

Note that support-minimal partition identities are primitive, since the circuits of a matrix are contained in the Graver basis of this matrix.

1.2.4 Integer programming and toric Gröbner bases

In this example you learn about the functions minimize, groebner, and normalform.

The following neat example is based on the example presented in [3]. Let us assume that we want to give change worth 99 cents using only pennies (1ct), nickels (5ct), dimes (10ct), and quarters (25ct). Clearly,

$$4 \cdot 1 + 4 \cdot 5 + 0 \cdot 10 + 3 \cdot 25 = 99$$

would be one way to do it. Is this there another choice of 11 coins that sums up to 99ct but uses fewer nickels and quarters (in total)? In other words, we would like to solve

$$\min\{x_2 + x_4 : x_1 + x_2 + x_3 + x_4 = 11, x_1 + 5x_2 + 10x_3 + 25x_4 = 99, x_1, x_2, x_3, x_4 \in \mathbb{Z}_+\}$$

Let us set up the problem in 4ti2.

4coins.mat	4coins.rhs	4coins.sign	4coins.cost
2 4	1 2	1 4	1 4
1 1 1 1	11 99	1 1 1 1	0 1 0 1
1 5 10 25			

Note that we do not have to specify a relations file 4coins.rel, since already by default all relations are assumed to be equations. Now we simply call

which creates the single output file

4coins.min						
1	4					
4	1	4	2			

From this, we conclude that

$$4 \cdot 1 + 1 \cdot 5 + 4 \cdot 10 + 2 \cdot 25 = 99$$

is an optimal choice, using only 3 instead of 7 nickels and quarters.

Remark. We could also specify a list of right-hand sides in 4coins.rhs. The call

./minimize 4coins

then creates a file 4coins.min containing minima to the corresponding integer programs.

Since we already know a feasible solution, there is another way we might attack this problem, namely via toric Gröbner bases. For this, we first need to specify the matrix A and the cost vector c in the two files accins.mat and accins.cost:

4coins.mat		4coins.cost							
2	4				1	4			
1	1	1	1		0	1	0	1	
1	5	10	25						

Then we compute the Gröbner basis of the toric ideal

$$I_A = \langle x^u - x^v : Au = Av, u, v \in \mathbb{Z}_+^4 \rangle$$

with respect to a term ordering \prec compatible with c, that is, $c^{\intercal}v < c^{\intercal}u$ implies $x^v \prec x^u$. This toric Gröbner basis is computed by

./groebner 4coins

and gives the output file

4coi:	ns.g	gro		
1	4			
4	4	0	3	

Then we specify our feasible solution in

4coins.zfeas							
1	4						
4	4	0	3				

and call

./normalform 4coins

to produce the file

that also contains the desired optimal solution.

Remark. We could also specify a list of feasible solutions in 4coins.zfeas. Then the call

./normalform 4coins

creates a file 4coins.nf containing the minima to the corresponding integer programs. (If z_0 is a feasible solution, the corresponding integer program is defined by putting the right-hand side to Az_0 .)

Rename 4coins.zfeas to 4coins.zfea?!

1.2.5 Markov Bases in Statistics

In this example you learn about the functions markov and genmodel.

Let us consider the following 4×4 table of non-negative integer numbers together with all row and column sums.

$$\begin{pmatrix}
11 & 23 & 34 & 3 \\
4 & 15 & 12 & 11 \\
17 & 2 & 3 & 25 \\
16 & 12 & 22 & 7
\end{pmatrix}$$

$$\begin{array}{c}
71 \\
42 \\
47 \\
57 \\
48 & 52 & 71 & 46
\end{array}$$

In statistics, one wishes to sample among arrays that have fixed counts, say fixed row and column sums. In order to sample, one needs a set of moves that, in particular, do not change the counts when added to the current table. Clearly, these moves must have counts 0 and thus quite naturally lead us to the toric ideal

$$I_A = \langle x^u - x^v : Au = Av, u, v \in \mathbb{Z}_+^{16} \rangle,$$

where

It turns out that for any set of fixed counts, a (minimal) Markov basis is given by a minimal generating set of this toric ideal. Note that a Markov basis connects all non-negative tables with these counts in the sense that for any two non-negative tables T_1 and T_2 with these counts, there is a sequence of non-negative tables $T_1 = S_0, \ldots, S_N = T_2$ with the same counts as T_1 and T_2 and such that $S_i - S_{i-1}$ or $S_{i-1} - S_i$ is in the Markov basis for $i = 1, \ldots, N$.

For two-way tables the situation is still very simple as our computations with 4×4 tables will now demonstrate. Write the matrix that defines our toric ideal in the file 4x4.mat:

4x4.	mat														
8	16														
1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0
0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1

Let us compute the Markov basis via the call

./markov 4x4

which creates a single output file 4x4.mar containing the 36 Markov basis elements. Up to symmetry (swapping rows or columns), the Markov basis consists of the single

move

$$\left(\begin{array}{cccc}
1 & -1 & 0 & 0 \\
-1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right).$$

In fact, this elementary move is (up to symmetry) the only representative of the minimal Markov moves for arbitrary $m \times n$ tables using row and column counts.

Creating the matrices for statistical models may be pretty cumbersome. 4ti2 provides a little function, genmodel, that helps the user with creating matrices for hierarchical models defined by a complex.

The $m \times n$ tables problem above corresponds to the complex $\{\{1\}, \{2\}\}$ on two nodes with levels m and n, respectively. Let us encode the complex for 3x6 tables with 1-marginals (row and column sums) in 3x6.mod.

3x6.mod							
3							
3	6						
2							
1	1						
1	2						

and call

./genmodel 3x6

to produce the desired matrix file 3x6.mat.

The encoding of the complex should be obvious from the example: first we state the number of nodes and their levels, then we give the number of maximal faces. Finally, we list each maximal face by first specifying the number of nodes on it and then by listing these nodes.

Thus, a 3x4x6 table with 2-marginals (that is, again only counts along coordinate axes) corresponds to the complex $\{\{(1,2)\},\{(2,3)\},\{(3,1)\}\}\}$ on 3 nodes with levels 3, 4, and 6, respectively. Thus, its encoding is in 4ti2 would look like:

3x4x6.mod							
3							
3	4	6					
3							
2	1	2					
2	2	3					
2	3	1					

A binary model on the bipartite graph $K_{2,3}$ then reads as

3x4x6.mod								
5								
2	2	2	2	2				
6								
2	1	3						
2	1	4						
2	1	5						
2	2	3						
2	2	4						
2	2	5						

Bibliography

- [1] Utz-Uwe Haus, Matthias Köppe, and Robert Weismantel. A primal all-integer algorithm based on irreducible solutions. *Math. Programming, Series B*, 96(2):205–246, 2003.
- [2] Matthias Köppe. Erzeugende Mengen für gemischt-Otto-von-Guerickeganzzahlige Programme. Diploma thesis, 1999. Universität Magdeburg, available from URL http://www.math.uni-magdeburg.de/~mkoeppe/art/mkoeppe-diplom.ps.
- [3] Bernd Sturmfels. Algebraic recipes for integer programming. arXiv:math.OC/0310194, 2003.
- [4] Regina Urbaniak. Decomposition of Generating Sets and of Integer Programs. Dissertation, Technische Universität Berlin, 1998.