MAST90053 EXPERIMENTAL MATHEMATICS

LECTURER: DR ANDREA BEDINI

2014, SEMESTER 1, WEEK 2

Interger relation detection

Given a real vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, an integer relation for x is a non-zero vector of integers $m = (m_1, m_2, \ldots, m_n) \in \mathbb{Z}^n$, such that $mx^t = m_1x_1 + m_2x_2 + \cdots + m_nx_n = 0$. An integer dectection algorithm is a computational scheme that can recover the vector of integers m (if it exists) or produce bounds within which no integer relation exists.

Although the integer relation problem is often regarded to be a relatively "new" problem, it is really a rather old problem. The problem of finding integer relations for two numbers (x_1, x_2) can be solved by applying the Euclidean algorithm to x_1, x_2, or, equivalently, by computing the continued fraction expansion of the real number x_1/x_2.

Integer relation detection methods are employed very often in experimental math applications to recognize a mathematical constant whose numerical value can be computed to at least moderately high precision, and also to discover relations between a set of computed numerical values.

The PSLQ algorithm

At the present time, the best known integer relation algorithm is the PSLQ algorithm by Ferguson, Bailey, and Arno (1999). The PSLQ algorithm, together with related lattice reduction schemes, was recently named one of ten "algorithms of the century" by the publication Computing in Science and Engineering (see Bailey 2000). In addition to possessing good numerical stability, PSLQ is guaranteed to find a relation in a polynomially bounded number of iterations. The name "PSLQ" derives from its usage of a partial sum of squares vector and a LQ (lower trapezoidal-orthogonal) matrix factorization. You can find all the proofs in the orignal paper by Ferguson, Bailey, and Arno (1999), on in Borwein (2002 app. B).

Definitions

Let $x \in \mathbb{R}^n$ be a nonzero n-tuple $x = (x_1, x_2, \ldots, x_n)$. Define the partial sum of squares, s_j, for x as

$$s_k = \sqrt{x_k^2 + \ldots + x_n^2}.$$

Naturally we can assume that x is a unit vector, so that $s_1 = |x| = 1$. Define the lower trapezoidal $n \times (n-1)$ matrix H_x with elements

$$h_{i,j} = \begin{cases} 0 & i < j \\ \frac{s_{i+1}}{s_i} & i = j \\ -\frac{x_i x_j}{s_j s_{j+1}} & i > j. \end{cases}$$

EXAMPLE 1. For $n = 4$, the matrix H_x is given by

$$H_x = \begin{pmatrix} \frac{s_2}{s_1} & 0 & 0 \\ -\frac{x_2 x_1}{s_1 s_2} & \frac{s_3}{s_2} & 0 \\ -\frac{x_3 x_1}{s_1 s_2} & -\frac{x_3 x_2}{s_2 s_3} & \frac{s_4}{s_3} \\ -\frac{x_4 x_1}{s_1 s_2} & -\frac{x_4 x_2}{s_2 s_3} & -\frac{x_4 x_3}{s_3 s_4} \end{pmatrix}$$

LEMMA 1. Let x be a unit vector and H_x the lower trapezoidal matrix defined above. Then:

1. Each column of H_x is orthogonal to x, i.e. $xH_x = 0$.
2. The columns of H_x form an orthonormal basis, i.e. $H_x^t H_x = I_{n-1}$.
3. $|H_x| = \sqrt{n-1}$, where $|A|$ is the Frobenius norm $|A| = (\sum_{i,j} a_{i,j}^2)^{1/2}$.

LEMMA 2. Let P_x be the $n \times n$ matrix given by $P_x = H_x H_x^t$. Then P_x satistfies the following:

1. $P_x^t = P_x$
2. $P_x = I_n - x^t x$
3. $P_x^2 = P_x$
4. $|P_x| = \sqrt{n-1}$
5. $P_x m^t = m^t$ for any relation m for x.

Hermite reduction

Suppose we can find an invertible $n \times n$ lower triangular *integer* matrix A such that a column of AH_x, say the jth one, has zeros everywhere except in the jth entry. Because we know that

$$0 = xH_x = xA^{-1}AH_x,$$

then the jth entry of the vector xA^{-1} is zero too. Up to an overall factor A^{-1} is also an integer matrix and therefore the above implies an integer relation for x whose entries are given by the jth column of A^{-1}.

If the matrix A were allowed to have real, rather than integer, entries, the above problem would be equivalent to solving a linear system and the solution would be obtained by reducing H in the row echelon form.

LEMMA 3. (Row echelon form). Let $H = (h_{i,j})$ be a $n \times (n-1)$ lower trapezoidal matrix of with $h_{ii} \neq 0$. Define an associated $n \times n$ lower triangular matrix $A = (a_{i,j})$ recursively as follows. For fixed i, decrement j from n to 1, setting

$$a_{i,j} = \begin{cases} 0 & j > i, \\ 1 & j = i, \\ -\sum_{j<k\leq i} a_{i,k} h_{k,j}/h_{j,j} & j < i. \end{cases}$$

Then AH is in the row echelon form:

$$AH = \begin{pmatrix} h_{1,1} & & & \\ & \ddots & & \\ & & & h_{n-1,n-1} \\ 0 & \cdots & & 0 \end{pmatrix}.$$

The analogue of the row echelon form for matrices over the integers $\mathbb{Z}$ is the Hermite normal form.

DEFINITION 1. (Hermite normal form). Let $H = (h_{i,j})$ be a $n \times (n-1)$ lower trapezoidal matrix of with $h_{ii} \neq 0$. Define an associated $n \times n$ lower triangular matrix $D = (d_{i,j})$ recursively as follows. For fixed i, decrement j from n to 1, setting

$$d_{i,j} = \begin{cases} 0 & j > i, \\ 1 & j = i, \\ -\operatorname{nint}\left(\sum_{j<k\leq i} d_{i,k} h_{k,j}/h_{j,j}\right) & j < i. \end{cases}$$

Then we will say $H' = DH$ is the *Hermite reduction* of H and that D is the *reducing matrix* of H. The function nint denotes the nearest integer function, e.g., $\operatorname{nint}(t) = \lfloor t + 1/2 \rfloor$.

DEFINITION 2. (Modified Hermite reduction). Let H be as above. We can obtain the reducing matrix D and its inverse $E = D^{-1}$ simultaneously with DH by the following procedure.

1. Set $D = E = I_n$.
2. Let H_i be the ith row of H, D_i the ith row of D and E_i the ith column of E.
3. For i from 2 to n and for j from $i-1$ to 1 (step -1)
   1. Set $q = \operatorname{nint}(h_{i,j}/h_{j,j})$
   2. Replace H_i by $H_i - qH_j$
   3. Replace D_i by $D_i - qD_j$
   4. Replace E_j by $E_j + qE_i$
4. H will be replaced by DH and $E = D^{-1}$.

Due to the nint function the off-diagonal elements of DH are not zero but it is easy to see that they are not too big either.

LEMMA 4. The entries of the Hermite reduced matrix $H' = (h'_{i,j}) = DH$ satisfy the inequality

$$|h'_{i,j}| \leq \frac{1}{2}|h'_{j,j}| \equiv \frac{1}{2}|h_{j,j}|.$$

This follows from the definitions of the nint function, Hermite reduction and the fact that $|t - \text{nint}(t)| \leq 1/2$.

PSLQ operates by constructing a series of matrices A_k, such that the entries of the vector $x_k = xA_k^{-1}$ steadily decrease in size. At any given iteration, the largest and smallest entries of x_k usually normally differ by no more than two or three orders of magnitude. When a relation is detected by the algorithm, the smallest entry of the x_k vector abruptly decreases to roughly the "epsilon" of the working precision (i.e. 10^{-p}, where p is the precision level in digits), and the desired relation is given by the corresponding column of A_k^{-1}.

Exchange and corner steps

A complementary way to understand PSLQ is given by the following theorem.

THEOREM 1. Let A be an invertible $n \times n$ matrix with integer coefficients and Q an orthogonal matrix $Q \in O(n-1)$. Let $x \in \mathbb{R}^n$, and H_x has before. If $H' = AH_xQ$ is lower trapezoidal with non-zero diagonal entries $h'_{j,j}$, then, for any integer relation m for x, we have:

$$\min_{1 \leq j < n} \frac{1}{|h'_{j,j}|} \leq |m|.$$

Therefore we can obtain a rising lower bound on the norm of m by reducing the diagonal of H. This is done by the exchange and corner steps.

Fix a real number $\gamma > 2/\sqrt{3}$, whose meaning will be explained later, and choose an integer r such that $\gamma^r |h_{r,r}|$ is maximal. Define the permutation matrix S_r to be the identity matrix with the r and $r+1$ rows exchanged and replace H with S_rH. This amounts to exchange the rows r and $r+1$ of H.

If $r < n-1$ then H is no longer trapezoidal but we can remedy with a simple rotation. Consider the 2×2 submatrix of H obtained taking the rows and columns r and $r+1$.

$$\begin{pmatrix} \alpha & 0 \\ \beta & \lambda \end{pmatrix} \equiv \begin{pmatrix} h_{r,r} & 0 \\ h_{r+1,r} & h_{r+1,r+1} \end{pmatrix}$$

Let $\delta = \sqrt{\beta^2 + \lambda^2}$. We can write

$$\begin{pmatrix} \delta & 0 \\ \alpha\beta/\delta & -\alpha\lambda/\delta \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ \beta & \lambda \end{pmatrix} \begin{pmatrix} \beta/\delta & -\lambda/\delta \\ \lambda/\delta & \beta/\delta \end{pmatrix}$$

where the last matrix is orthogonal.

Given our choice of r we have $|h_{r+1,r+1}| \leq \frac{1}{\gamma}|h_{r,r}|$. Combining this fact with $|h_{r,r+1}| \leq 1/2|h_{r,r}|$ from Lemma 4 we can see that the diagonal element $h_{r,r}$ is reduced as long as

$$\sqrt{\frac{1}{4} + \frac{1}{\gamma^2}} < 1 \text{ or } \gamma > 2/\sqrt{3}.$$

More formally we can defined the $(n-1) \times (n-1)$ orthogonal matrix Q_r as follows. If $r = n-1$ set $Q_{n-1} = I_{n-1}$. Otherwise, for $1 \leq r < n-1$, the entries of Q_r are given by,

$$\begin{aligned} q_{i,i} &= 1 && \text{for } i \neq r, r+1 \\ \begin{pmatrix} q_{r,r} & q_{r,r+1} \\ q_{r+1,r} & q_{r+1,r+1} \end{pmatrix} &= \begin{pmatrix} \beta/\delta & -\lambda/\delta \\ \lambda/\delta & \beta/\delta \end{pmatrix} \\ q_{i,j} &= 0 && \text{otherwise} \end{aligned}$$

With this notation the corner step can be defined as replacing $S_r H$ from the exchange step with $S_r H Q_r$.

Statement of the algorithm

We now want to put all the pieces together and formulate the algorithm. The algorithm takes a input vector $x \in \mathbb{R}^n$, known with a precision of p digits, a constant $\gamma > 2/\sqrt{3}$ and an optional target bound M^* on the norm of any possible relation for x.

1. Set $H = H_x$ as defined above and set the matrices $A = B = I_n$.
2. Compute the Hermite reducing matrix D. Replace H by DH, A by DA and B by BD^{-1}, x by xD^{-1}.
3. Update the lower bound on $|m| \geq M = \min_{1 \leq j < n} |h_{j,j}|^{-1}$. Optionally, terminate the algorithm if $M > M^*$.

4. Choose r such that $\gamma^r |h_{r,r}|$ is maximal.
5. Replace H by $S_r H Q_r$, A by $S_r A$, B by $B S_r$, x by $x S_r$.
6. Perform the reduction as in step 2.
7. Terminate the algorithm if $x_j < \epsilon$ or $h_{j,j} < \epsilon$ for some j, where ϵ is small quantity close to the working floating point precision e.g. 10^{-p+2}.

At the end of this iteration either a integer relation m for x will appear as the column of B corresponding to the zero entry of x or the algorithm has determined that no relation exist of norm $|m| \leq M^*$.

References

Bailey, David H. 2000. "Integer Relation Detection." *Computing in Science & Engineering* 2 (1). AIP Publishing: 24–28.

Borwein, Peter. 2002. *Computational Excursions in Analysis and Number Theory*. Vol. 10. Springer.

Ferguson, Helaman, David H Bailey, and Steve Arno. 1999. "Analysis of PSLQ, an Integer Relation Finding Algorithm." *Mathematics of Computation of the American Mathematical Society* 68 (225): 351–69. http://www.ams.org/journals/mcom/1999-68-225/S0025-5718-99-00995-3/S0025-5718-99-00995-3.pdf.