Chió's Trick for Linear Equations with Integer Coefficients

Gaussian Elimination solves a system $Ax = b$ of linear equations for x by a sequence of rational operations $(+, -, \cdot, /)$ during which rounding errors occur unless some extra effort is put into performing the arithmetic exactly. This effort may be worthwhile if the data $[A, b]$ is known exactly, particularly if the data consists entirely of integers. In this case the arithmetic generates rational numbers stored as pairs of integers " in lowest terms " (*i.e.* with no common divisor). Reduction to lowest terms is time-consuming but necessary to prevent the pairs of integers from growing enormously too wide, and it reveals a curious phenomenon: the divisor integers take relatively few distinct values. Chió's trick exploits this phenomenon to perform elimination using exclusively integer arithmetic with integers scarcely bigger than they have to be.

Since its appearance in 1853 Chió's trick has been rediscovered repeatedly (once by the Rev. C.L. Dodgson who wrote *Alice in Wonderland* and other amusements under his pen-name "Lewis Carroll") and often with incomplete, incorrect or extremely complicated proofs of validity. Let's see whether these notes can do better.

Gaussian Elimination as Triangular Factorization

Nowadays many texts explain the relation between Gaussian Elimination and the Triangular Factorization $PA = LU$ wherein P is a permutation matrix that takes account of "Pivotal" row exchanges, L is "Unit Lower Triangular" (it has 1 's on its diagonal), and U is upper triangular. Although P is a product of row exchanges performed as they are determined during the elimination process, we can imagine that they had been applied in advance to produce a matrix PA whose linearly dependent rows, if any, are its last. It is customary to take for granted also that the linearly dependent columns of PA , if any, are its last few, as can be arranged by reordering columns if necessary. In short, all except perhaps the last few of the leading principal submatrices of PA are assumed invertible; otherwise the factorization $PA = LU$ becomes either impossible or non-unique.

Our treatment of Chió's trick is also simplified by the assumption that all except perhaps the last few of the leading principal submatrices of A are invertible. This simplification is tantamount to the application in advance of whatever row and/or column exchanges would otherwise be found during Chió's elimination process to be necessary to ensure the assumption's validity; this is not a significant restriction so far as matrix computations are concerned. However, this restriction does prevent our treatment of Chió's trick from being applied directly to explain how a similar trick works during the computations of greatest common divisors of pairs of polynomials, and the computations of continued fraction expansions of rational functions.

Our treatment begins with a representation for the intermediate stages reached during the process of Gaussian elimination or, equivalently, triangular factorization. For each $k = 1, 2, 3, \dots$, after the first k unknowns kave been eliminated from all but the first k of the equations $Ax = b$, the remaining equations take the form $S_k z_k = g_k$ wherein z_k is obtained from x by deleting its first k components, and S_k is a *Schur Complement* derived from A by partitioning as follows:

$$[A, b] = \begin{bmatrix} V_k & R_k & r_k \\ C_k & H_k & h_k \end{bmatrix} = \begin{bmatrix} L_k & O \\ \bar{L}_k & I \end{bmatrix} \cdot \begin{bmatrix} U_k & \bar{U}_k & \bar{g}_k \\ O & S_k & g_k \end{bmatrix} .$$

Here $V_k = L_kU_k$ is the first principal k-by-k submatrix of A ; its triangular factors are L_k and U_k . The existence of V_k^{-1} is assured by our simplifying assumption, which also ensures the existence of the factorization $V_k = L_kU_k$. Because L_k is unit triangular, $\det(L_k) = 1$ and we find $\det(U_k) = \det(L_k)\det(U_k) = \det(L_kU_k) = \det(V_k) \neq 0$, so U_k^{-1} exists and $U_k^{-1}L_k^{-1} = V_k^{-1}$. Also $\bar{L}_k := C_kU_k^{-1}$ and $[\bar{U}_k, \bar{g}_k] := L_k^{-1}[R_k, r_k]$. The Schur complement of V_k in A is S_k determined from $[S_k, g_k] := [H_k, h_k] - C_kV_k^{-1}[R_k, r_k]$. (Check it all out!)

What happens when k advances to $k+1$? L_k becomes the leading $(k+1)$-by-$(k+1)$ principal submatrix of L_{k+1} , and U_k does likewise for U_{k+1} . The rest is best explained by partitioning

$$[S_k, g_k] = \begin{bmatrix} \beta_k & p^T{}_k & \pi_k \\ q_k & W_k & w_k \end{bmatrix} .$$

First $ß_k$ becomes the last element $u_{k+1,k+1}$ of U_{k+1} ; the rest of the last column of U_{k+1} comes from the first column of $\bar{U}_k$. The rest of $[\bar{U}_k, \bar{g}_k]$ becomes the first k rows of $[\bar{U}_{k+1}, \bar{g}_{k+1}]$, whose last row comes from $[p^T{}_k, \pi_k]$. The last row of L_{k+1} is formed by appending 1 to the first row of $\bar{L}_k$, whose remaining rows form the first k columns of $\bar{L}_{k+1}$, whose last column is $q_k/ß_k$. Finally another pass of elimination produces the recurrence

$$[S_{k+1}, g_{k+1}] := [W_k, w_k] - (q_k/ß_k)[p^T{}_k, \pi_k] .$$

VERIFY THE FOREGOING PARAGRAPH TO CONFIRM YOUR UNDERSTANDING
OF THE PROCESSES OF ELIMINATION AND TRIANGULAR FACTORIZATION

For future reference note that $ß_k = u_{k+1,k+1} = \det(U_{k+1})/\det(U_k) = \det(V_{k+1})/\det(V_k)$.

Since $V_k^{-1} = \mathrm{Adj}(V_k)/\det(V_k)$, we see that $S_k = H_k - C_k\,\mathrm{Adj}(V_k)\,R_k/\det(V_k)$ is a rational function of the elements of A with common† denominator $\det(V_k)$. Therefore all elements of

$$[T_k, u_k] := \det(V_k)[S_k, g_k] = \det(V_k)[H_k, h_k] - C_k\,\mathrm{Adj}(V_k)\,[R_k, r_k]$$

are polynomials in the elements of A , and the reduced equations $S_kz_k = g_k$ are equivalent to equations $T_kz_k = u_k$ all of whose coefficients are, like those of $[A, b]$, integers.

This is what Chió's trick does, but not directly. If we tried to compute polynomials T_k and u_k directly using only additions, subtractions and multiplications, but no divisions, the arithmetic work would grow horrendously with k . Chió's trick works faster by using divisions too.

. . . .

† Footnote: Though $V_k^{-1} = \mathrm{Adj}(V_k)/\det(V_k)$ and $S_k = H_k - C_k\,\mathrm{Adj}(V_k)\,R_k/\det(V_k)$ are rational functions " of the elements of A with common denominator $\det(V_k)$ " some elements of V_k^{-1} and S_k may, after reduction to lowest terms, have denominators that properly divide $\det(V_k)$. This certainly happens when V_k is triangular, for instance.

Chió’s Trick

It is an algorithm that, for each $k = 1, 2, 3, \dots$ in turn, shall be shown to compute the coefficients $[T_k, u_k]$ of a set of equations $T_k z_k = u_k$ with the same solution z_k as the reduced set $S_k z_k = g_k$ obtained from Gaussian Elimination, but the elements of $T_k := \det(V_k) S_k$ are polynomials in the elements of A instead of rational functions like the Schur complement $S_k = H_k - C_k V_k^{-1} R_k$. Let us disregard the right-hand side columns $b, h_k, r_k, \bar{g}_k, g_k, u_k$ for the time being since they’re just along for the ride.

Chió’s algorithm defines a sequence of matrices $A_{(k)}$ with elements $a_{(k)ij}$ thus:

$a_{(0)00} := \mu_{(0)} := 1$;

$A_{(1)} := A$ so $a_{(1)ij} := a_{ij}$ for all $i > 0$ and $j > 0$;

for $k = 1, 2, 3, \dots$ in turn,

$$a_{(k+1)ij} := (\, a_{(k)kk} \cdot a_{(k)ij} - a_{(k)ik} \cdot a_{(k)kj} \,)/a_{(k-1)(k-1)(k-1)} \quad \text{for all } i > k \text{ and } j > k .$$

Of course, the algorithm would fail if any $a_{(k-1)(k-1)(k-1)} = 0$, so this will have to be proved impossible because of our simplifying assumption about invertible leading principal submatrices.

Partitioning $A_{(k)} = \begin{bmatrix} \mu_{(k)} & m^T_{(k)} \\ f_{(k)} & M_{(k)} \end{bmatrix}$ provides a compact description of Chió’s algorithm:

$$A_{(k+1)} := (\, \mu_{(k)} M_{(k)} - f_{(k)} m^T_{(k)} \,)/\mu_{(k-1)} .$$

Our inductive proof of its effectiveness begins with the induction hypotheses that $\mu_{(k)} = \det(V_k)$ and that $A_{(k)} = \mu_{(k-1)} S_{k-1} = T_{k-1}$, *i.e.*, that

$$\begin{bmatrix} \mu_{(k)} & m^T_{(k)} \\ f_{(k)} & M_{(k)} \end{bmatrix} = \mu_{(k-1)} \begin{bmatrix} \beta_{k-1} & p^T_{k-1} \\ q_{k-1} & W_{k-1} \end{bmatrix} .$$

These hold at $k = 1$ because $\mu_{(0)} = 1$ and $A_{(1)} = A = S_0$, so $\mu_{(1)} = \mu_{(0)} ß_0 = a_{11} = \det(V_1) \neq 0$. Now suppose the hypotheses hold for $k = 1, 2, \dots, K$ and recall from “ future reference ” that $ß_{K-1} = \det(V_K)/\det(V_{K-1}) = \mu_{(K)}/\mu_{(K-1)}$. Then

$$\begin{aligned} A_{(K+1)} &= (\, \mu_{(K)} M_{(K)} - f_{(K)} m^T_{(K)} \,)/\mu_{(K-1)} && \text{from Chió's algorithm,} \\ &= \mu_{(K-1)} (\, ß_{K-1} W_{K-1} - q_{K-1} p^T_{K-1} \,) && \text{from the second induction hypothesis,} \\ &= \mu_{(K)} \left(W_{K-1} - (q_{K-1}/ß_{K-1}) p^T_{K-1} \right) && \text{because } ß_{K-1} = \mu_{(K)}/\mu_{(K-1)} , \\ &= \mu_{(K)} S_K = T_K && \text{from the recurrence for } S_K \text{ and definition of } T_K . \end{aligned}$$

This confirms the second induction hypothesis for $k = K+1$, from which the observation that $\mu_{(K+1)} = \mu_{(K)} ß_K = \det(V_{K+1})$ confirms the first. End of proof.

To deal with the right-hand side columns $b, \dots, u_k$ just do unto $[A, b]$ whatever row operations Chió’s algorithm does unto A . Thus Chió’s algorithm reduces the given linear system $Ax = b$ through a sequence of ever smaller systems $T_k z_k = u_k$ whose elements are polynomials in the data $[A, b]$; in fact $[T_k, u_k] = \det(V_k)[H_k, h_k] - C_k \operatorname{Adj}(V_k)\, [R_k, r_k\,]$ is of total degree $k+1$. If the division by $\mu_{(k-1)}$ were omitted from Chió’s algorithm, the total degree would be 2^k instead. This is why, when k is big, Chió’s trick saves a lot of work during exact computation with integers or symbolic algebraic data.