Lecture Notes 1: Matrix Algebra
Part C: Pivoting and Matrix Decomposition

Peter J. Hammond

Autumn 2012, revised Autumn 2014

Lecture Outline

Outline

Triangular Matrices: Definition

Definition

A square matrix is upper (resp. lower) triangular
if all its non-zero off diagonal elements are above and to the right
(resp. below and to the left) of the diagonal
— i.e., in the upper (resp. lower) triangle
bounded by the principal diagonal.

- The elements of an upper triangular matrix $\mathbf{U}$ satisfy $(\mathbf{U})_{ij} = 0$ whenever $i > j$.
- The elements of a lower triangular matrix $\mathbf{L}$ satisfy $(\mathbf{L})_{ij} = 0$ whenever $i < j$.

Triangular Matrices: Exercises

Exercise

Prove that the transpose:

1. $\mathbf{U}^\top$ *of any upper triangular matrix* $\mathbf{U}$ *is lower triangular;*
2. $\mathbf{L}^\top$ *of any lower triangular matrix* $\mathbf{L}$ *is upper triangular.*

Exercise

Consider the matrix $\mathbf{E}_{r+\alpha q}$
that represents the elementary row operation
of adding a multiple of α *times row* q *to row* r.

Under what conditions is $\mathbf{E}_{r+\alpha q}$
(i) upper triangular? (ii) lower triangular?

Hint: Apply the row operation to the identity matrix $\mathbf{I}$.

Answer: (i) iff $q < r$; (ii) iff $q > r$.

Products of Upper Triangular Matrices

Theorem

The product $\mathbf{W} = \mathbf{UV}$ *of any two upper triangular matrices* $\mathbf{U}, \mathbf{V}$ *is upper triangular,*
with diagonal elements $w_{ii} = u_{ii} v_{ii}$ $(i = 1, \ldots, n)$ *equal to the product of the corresponding diagonal elements of* $\mathbf{U}, \mathbf{V}$.

Proof.

Given any two upper triangular $n \times n$ matrices $\mathbf{U}$ and $\mathbf{V}$, the elements $(w_{ij})^{n \times n}$ of their product $\mathbf{W} = \mathbf{UV}$ satisfy

$$w_{ij} = \begin{cases} \sum_{k=i}^{j} u_{ik} v_{kj} & \text{if } i \leq j \\ 0 & \text{if } i > j \end{cases}$$

because $u_{ik} v_{kj} = 0$ unless both $i \leq k$ and $k \leq j$.

So $\mathbf{W} = \mathbf{UV}$ is upper triangular.

Finally, putting $j = i$ implies that $w_{ii} = u_{ii} v_{ii}$ for $i = 1, \ldots, n$. □

Products of Lower Triangular Matrices

Theorem

The product of any two lower triangular matrices is lower triangular.

Proof.

Given any two lower triangular matrices $\mathbf{L}, \mathbf{M}$,
taking transposes shows that $(\mathbf{LM})^\top = \mathbf{M}^\top \mathbf{L}^\top = \mathbf{U}$,
where the product $\mathbf{U}$ is upper triangular,
as the product of upper triangular matrices.

Hence $\mathbf{LM} = \mathbf{U}^\top$ is lower triangular,
as the transpose of an upper triangular matrix. □

Determinants of Triangular Matrices

Theorem

The determinant of any $n \times n$ upper triangular matrix $\mathbf{U}$ equals the product of all the elements on its principal diagonal.

Proof.

Recall the expansion formula $|\mathbf{U}| = \sum_{\pi \in \Pi} \mathrm{sgn}(\pi) \prod_{i=1}^{n} u_{i\pi(i)}$ where Π denotes the set of permutations on $\{1, 2, \ldots, n\}$.

Because $\mathbf{U}$ is upper triangular, one has $u_{i\pi(i)} = 0$ unless $i \leq \pi(i)$.

So $\prod_{i=1}^{n} u_{i\pi(i)} = 0$ unless $i \leq \pi(i)$ for all $i = 1, 2, \ldots, n$.

But the only permutation $\pi \in \Pi$ which satisfies $i \leq \pi(i)$ for all $i = 1, 2, \ldots, n$ is the identity permutation ι.

Because $\mathrm{sgn}(\iota) = 1$, the expansion reduces to the single term

$$|\mathbf{U}| = \mathrm{sgn}(\iota) \prod_{i=1}^{n} u_{i\iota(i)} = \prod_{i=1}^{n} u_{ii}$$

which is the product of the diagonal elements, as claimed. $\square$

Inverting Triangular Matrices

Similarly $|\mathbf{L}| = \prod_{i=1}^{n} \ell_{ii}$ for any lower triangular matrix $\mathbf{L}$.
Evidently:

Corollary

A triangular matrix (upper or lower) is invertible if and only if no element on its principal diagonal is 0.

In the next slide, we shall prove:

Theorem

If the inverse $\mathbf{U}^{-1}$ of an upper triangular matrix $\mathbf{U}$ exists, then it is upper triangular.

Taking transposes leads immediately to:

Corollary

If the inverse $\mathbf{L}^{-1}$ of an lower triangular matrix $\mathbf{L}$ exists, then it is lower triangular.

Inverting Triangular Matrices: Proofs

Recall the $(n-1) \times (n-1)$ cofactor matrix $\mathbf{C}_{rs}$
that results from omitting row r and column s of $\mathbf{U} = (u_{ij})$.

When it exists, $\mathbf{U}^{-1} = (1/|\mathbf{U}|)\,\mathbf{adj}\,\mathbf{U}$, so it is enough to prove
that the $n \times n$ matrix $(|\mathbf{C}_{rs}|)$ of cofactor determinants,
whose transpose $(|\mathbf{C}_{rs}|)^\top$ is the adjugate, is lower triangular.

In case $r < s$, every element below the diagonal of the matrix $\mathbf{C}_{rs}$
is also below the diagonal of $\mathbf{U}$, so must equal 0.

Hence $\mathbf{C}_{rs}$ is upper triangular,
with determinant equal to the product of its diagonal elements.

Yet $s - r$ of these diagonal elements are $u_{i+1,i}$ for $i = r, \ldots, s-1$.
These elements are from below the diagonal of $\mathbf{U}$, so equal zero.

Hence $r < s$ implies that $|\mathbf{C}_{rs}| = 0$, so the $n \times n$ matrix $(|\mathbf{C}_{rs}|)$
of cofactor determinants is indeed lower triangular, as required. $\square$

Outline

Unitriangular Matrices: Definition and Two Properties

Definition

A unitriangular matrix is a triangular matrix (upper or lower) for which all elements on the principal diagonal equal 1.

Theorem

The determinant of any unitriangular matrix is 1.

Proof.

The determinant of any triangular matrix is the product of its diagonal elements, which must be 1 in the unitriangular case when every diagonal elements is 1. $\square$

Converting a Diagonal Matrix to Unitriangular Form

Theorem

Suppose $\mathbf{U}$ *is any upper triangular matrix*
with the property that all its diagonal elements $u_{ii} \neq 0$*.*

Then there exists a diagonal matrix $\mathbf{D}$
such that both $\mathbf{DU}$ *and* $\mathbf{UD}$ *are upper unitriangular.*

Similarly for any lower triangular matrix $\mathbf{L}$
with the property that all its diagonal elements $\ell_{ii} \neq 0$*.*

Converting a Diagonal Matrix: Proof

Define $\mathbf{D}$ as the diagonal matrix $\mathbf{diag}\,((1/u_{ii})_{i=1}^{n})$
whose diagonal elements d_{ii} are the reciprocals $1/u_{ii}$
of the corresponding elements u_{ii} of the upper triangular matrix $\mathbf{U}$,
all of which are assumed to be non-zero.

Then $\mathbf{DU}$ is upper unitriangular because $(\mathbf{DU})_{ik} = d_{ii}\delta_{ik}$ and so

$$(\mathbf{DU})_{ij} = \sum\nolimits_{k=1}^{n} d_{ii}\delta_{ik}u_{kj} = d_{ii}u_{ij} = \begin{cases} 1 & \text{when } i = j; \\ 0 & \text{when } i > j. \end{cases}$$

The same holds for $\mathbf{UD}$ whose elements $(\mathbf{UD})_{ij} = u_{ij}d_{jj}$
are also 1 when $i = j$ and 0 when $i > j$. $\square$

The Product of Unitriangular Matrices Is Unitriangular

Theorem

The product $\mathbf{W} = \mathbf{UV}$ *of any two upper unitriangular* $n \times n$ *matrices* $\mathbf{U}$ *and* $\mathbf{V}$ *is also upper unitriangular.*

Proof.

Because both $\mathbf{U}$ and $\mathbf{V}$ are upper triangular, so is $\mathbf{W} = \mathbf{UV}$.

Also, each i element of the principal diagonal of $\mathbf{W}$ is $w_{ii} = u_{ii} v_{ii}$, which is 1 because unitriangularity implies that $u_{ii} = v_{ii} = 1$.

It follows that $\mathbf{W}$ is upper unitriangular. $\square$

The same argument can be used to show that the product of any two lower unitriangular $n \times n$ matrices is also lower unitriangular.

The Inverse of a Unitriangular Matrix Is Unitriangular

Theorem

Any upper unitriangular $n \times n$ matrix $\mathbf{U}$ is invertible, with an upper unitriangular inverse $\mathbf{U}^{-1}$.

Proof.

Because $\mathbf{U}$ is unitriangular, its determinant is 1, so $\mathbf{V} = \mathbf{U}^{-1}$ exists.

Because $\mathbf{U}$ is upper triangular, so is $\mathbf{U}^{-1}$.

Also $u_{ii} v_{ii} = \delta_{ii} = 1$ for all $i = 1, 2, \ldots, n$, implying that $v_{ii} = 1/u_{ii} = 1$.

Therefore $\mathbf{U}^{-1}$ is indeed upper unitriangular. □

The same argument can be used to show that the inverse of any lower unitriangular $n \times n$ matrix is also lower unitriangular.

Outline

Three Simultaneous Equations

Consider the system
of three simultaneous equations in three unknowns,
which depends upon two "exogenous" constants a and b:

$$\begin{array}{ccccccc} x & + & y & - & z & = & 1 \\ x & - & y & + & 2z & = & 2 \\ x & + & 2y & + & az & = & b \end{array}$$

It can be expressed as using an augmented 3×4 matrix:

$$\begin{array}{rrr|r} 1 & 1 & -1 & 1 \\ 1 & -1 & 2 & 2 \\ 1 & 2 & a & b \end{array}$$

or, perhaps more usefully, a doubly augmented 3×7 matrix:

$$\begin{array}{rrr|r|rrr} 1 & 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 2 & a & b & 0 & 0 & 1 \end{array}$$

whose last 3 columns are those of the 3×3 identity matrix $\mathbf{I}_3$.

The First Pivot Step

Start with the doubly augmented 3×7 matrix:

$$\begin{array}{ccc|c|ccc} 1 & 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 2 & a & b & 0 & 0 & 1 \end{array}$$

First, we pivot about the element in row 1 and column 1 to zeroize the other elements of column 1.

This elementary row operation requires us to subtract row 1 from both rows 2 and 3. It is equivalent to multiplying by the lower triangular matrix $\mathbf{E}_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$.

Note that is the result of applying the same row operation to $\mathbf{I}$.

The result is:

$$\begin{array}{ccc|c|ccc} 1 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 3 & 2 & -1 & 1 & 0 \\ 0 & 1 & a+1 & b-1 & -1 & 0 & 1 \end{array}$$

The Second Pivot Step

After augmenting again by the identity matrix, we have:

$$\begin{array}{rrr|r|rrr|rrr} 1 & 1 & -1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 3 & 2 & -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & a+1 & b-1 & -1 & 0 & 1 & 0 & 0 & 1 \end{array}$$

Next, we pivot about the element in row 2 and column 2.
Specifically, multiply the second row by $-\frac{1}{2}$,
then subtract the new second row from the third to obtain:

$$\begin{array}{rrr|r|rrr|rrr} 1 & 1 & -1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & a+\frac{5}{2} & b-\frac{1}{2} & -\frac{3}{2} & \frac{1}{2} & 1 & 0 & \frac{1}{2} & 1 \end{array}$$

Again, the pivot operation is equivalent to multiplying

by the lower triangular matrix $\mathbf{E}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$,

which is the result of applying the same row operation to $\mathbf{I}$.

Case 1: Dependent Equations

In case 1, when $a + \frac{5}{2} = 0$, the equation system reduces to:

$$\begin{array}{rcrcrcr} x & + & y & - & z & = & 1 \\ & & y & - & \frac{3}{2}z & = & -\frac{1}{2} \\ & & & & 0 & = & b - \frac{1}{2} \end{array}$$

In case 1A, when $b \neq \frac{1}{2}$, neither the last equation, nor the system as a whole, has any solution.

In case 1B, when $b = \frac{1}{2}$, the third equation is redundant.

The first two equations have a general solution with $y = \frac{3}{2}z - \frac{1}{2}$ and $x = z + 1 - y = \frac{3}{2} - \frac{1}{2}z$, where z is arbitrary.

In particular, there is an entire one-dimensional space of solutions.

Case 2: Three Independent Equations

$$\begin{array}{ccc|c|ccc|ccc}
1 & 1 & -1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 \\
0 & 0 & a+\frac{5}{2} & b-\frac{1}{2} & -\frac{3}{2} & \frac{1}{2} & 1 & 0 & -\frac{1}{2} & 1
\end{array}$$

Case 2 occurs when $a + \frac{5}{2} \neq 0$,
and so the reciprocal $c := 1/(a + \frac{5}{2})$ is well defined.

Now divide the last row by $a + \frac{5}{2}$, or multiply by c, to obtain:

$$\begin{array}{ccc|c|ccc|}
1 & 1 & -1 & 1 & 1 & 0 & 0 \\
0 & 1 & -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 \\
0 & 0 & 1 & (b-\frac{1}{2})c & -\frac{3}{2}c & \frac{3}{2}c & \frac{1}{2}c
\end{array}$$

The system has been reduced to row echelon form in which:

1. the leading non-zero element of each row equals 1;
2. the leading zeroes of each row form the steps of a ladder (or *échelle*) which descends as one goes from left to right.

Case 2: Three Independent Equations, Third Pivot

$$\begin{array}{rrr|r|rrr|} 1 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & (b-\frac{1}{2})c & -\frac{3}{2}c & \frac{3}{2}c & \frac{1}{2}c \end{array}$$

Next, we zeroize the elements in the third column above row 3.

To do so, pivot about the element in row 3 and column 3.

This requires adding the last row to the first,
and $\frac{3}{2}$ times the last row to the second.

In effect, one multiplies

by the upper triangular matrix $\mathbf{E}_3 := \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix}$

The first three columns of the result are $\begin{array}{ccc} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}$

Case 2: Three Independent Equations, Final Pivot

$$\begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

The final pivoting operation involves subtracting the second row from the first, so the first three columns become the identity matrix

$$\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

This is a matrix in reduced row echelon form because, given the leading non-zero element of any row (if there is one), all elements above this element are zero.

Final Exercise

Exercise

1. *Find the last 4 columns of each* 3×7 *matrix produced by these last two pivoting steps.*
2. *Check that the fourth column solves the original system of 3 simultaneous equations.*
3. *Check that the last 3 columns form the inverse of the original coefficient matrix.*

Outline

Definition

An $m \times n$ matrix is in row echelon form just in case:

1. The first non-zero element in each row, called the leading entry, is 1.

   That is, in each row $r \in \{1, 2, \dots, m\}$, there is leading element $a_{r\ell}$ for which:

   - $a_{r\ell} = 1$;
   - $a_{rc} = 0$ for all $c < \ell$.

2. Each leading entry is in a column to the right of the leading entry in the previous row.

   This requires that, given the leading element $a_{r\ell} = 1$ of row r, one has $a_{r'c} = 0$ for all $r' > r$ and all $c \leq \ell$.

3. In case a row has no leading entry, because all its elements are zero, it must be below any row with a leading entry.

Examples

Here are three examples of matrices in row echelon form

$$\mathbf{A}_{\text{ref}} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ; \; \mathbf{B}_{\text{ref}} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} ; \; \mathbf{C}_{\text{ref}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Here are three examples of matrices
that are not in row echelon form

$$\mathbf{D} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} ; \; \mathbf{E} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} ; \; \mathbf{F} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

A Generalized Row Echelon Form

An $m \times n$ matrix is in generalized row echelon form (or GREF) just in case:

1. Each (non-zero) leading entry is in a column to the right of the leading entry in the previous row.

   This requires that, given the leading element $a_{r\ell} \neq 0$ of row r, one has $a_{r'c} = 0$ for all $r' > r$ and all $c \leq \ell$.
2. In case a row has no leading entry,
   because all its elements are zero,
   it must be below any row with a leading entry.

That is, we abandon the restriction
that the first non-zero element in each row is 1.

Pivoting to Reach a Generalized Row Echelon Form

Any $m \times n$ matrix can be transformed into its row echelon form by applying a series of elementary row operations involving non-zero pivot elements.

1. Look for the first non-zero column j_1 in the matrix, and find within it an element $a_{i_1 j_1} \neq 0$ with a large absolute value $|a_{i_1 j_1}|$; this will be the first pivot.
2. Interchange rows 1 and i_1, moving the pivot to the top row.
3. Subtract a_{ij_1}/a_{1j_1} times the new row 1 from each new row $i > 1$.

   This first pivot operation will zeroize all the elements of the pivot column j_1 that lie below the new row 1.

The Intermediate Matrices and Pivot Steps

After $k - 1$ pivoting operations have been completed,
and column j_{k-1} (with $j_{k-1} \geq k - 1$) was the last to be used:

1. The first $k - 1$ rows of the $m \times n$ matrix form a $(k - 1) \times n$ GREF matrix.
2. The last $m - k + 1$ rows of the $m \times n$ matrix form an $(m - k + 1) \times n$ matrix whose first j_{k-1} columns are all zero.
3. To determine the next pivot, look for the first column j_k which has a non-zero element below row $k - 1$, and find within it an element $a_{i_k j_k} \neq 0$ with $i_k \geq k$ and with a large absolute value $|a_{i_k j_k}|$; this will be the kth pivot.
4. Interchange rows k and i_k, moving the pivot up to row k.
5. Subtract a_{ij_k}/a_{kj_k} times the new row k from each new row $i > k$.

   This kth pivot operation will zeroize all the elements of the pivot column j_k that lie below the new row k.

Ending the Pivoting Process

1. Continue pivoting about successive pivot elements $a_{i_k j_k} \neq 0$, moving row $i_k \geq k$ up to row k at each stage k, while leaving all rows above k unchanged.
2. Stop after r steps when either $r = m$, or else all elements in the remaining $m - r$ rows are zero, so no further pivoting is possible.

Permuting before Pivoting

Suppose that pivoting stops after r steps.

Suppose that the elements $(a_{i_k j_k})_{k=1}^r$
of the original $m \times n$ matrix $\mathbf{A}$ have been used as the r pivots.

Let $\mathbf{P}$ denote the $m \times m$ permutation matrix whose kth row
satisfies $\mathbf{p}_k^\top = (p_{kj})_{j=1}^n = (\delta_{i_k j})_{j=1}^n$ for all $k \in \mathbb{N}_r$,
so that each row k of $\mathbf{P}$ equals row i_k of the identity matrix $\mathbf{I}_m$.

Also, in case pivoting stops with $r < m$,
suppose that rows $r+1, \ldots, m$ of $\mathbf{P}$
are chosen arbitrarily from non-pivot rows of $\mathbf{A}$.

Then the elements of the $m \times n$ matrix $\mathbf{PA}$ satisfy

$$(\mathbf{PA})_{kj} = \sum\nolimits_{\ell=1}^{m} p_{k\ell} a_{\ell j} = \sum\nolimits_{\ell=1}^{m} \delta_{i_k \ell} a_{\ell j} = a_{i_k j}$$

Pivoting after Permuting

Then the $m \times n$ matrix $\tilde{\mathbf{A}} := \mathbf{PA}$ that results from these operations can be transformed to GREF form by pivoting successively about its elements $(\tilde{a}_{kj_k})_{k=1}^r$.

Remember that the kth pivoting operation involves subtracting a multiple $\tilde{a}_{ij_k}/\tilde{a}_{kj_k}$ of the pivot row k from each lower row i (with $i > k$), in order to zeroize the ij_k element for all $i > k$.

For each $k \in \mathbb{N}_r$, the kth pivoting operation is therefore represented by a lower unitriangular $m \times m$ matrix $\tilde{\mathbf{L}}_k$.

So then is the product matrix $\mathbf{L} := \tilde{\mathbf{L}}_r \tilde{\mathbf{L}}_{r-1} \ldots \tilde{\mathbf{L}}_2 \tilde{\mathbf{L}}_1$ that results from combining all the successive pivoting operations into a single transformation.

Hence, there exists an $m \times m$ lower unitriangular matrix $\mathbf{L}$ such that $\mathbf{LPA}$ is in GREF.

Outline

Definitions

A matrix is in reduced row echelon form (RREF)
(respectively, in generalized reduced row echelon form (GRREF))
when it satisfies the following conditions.

1. The matrix is in row echelon form
(respectively, in generalized row echelon form).
2. The leading entry $a_{i\ell} \neq 0$ in each row i
is the only non-zero entry in its column.

   That is, $a_{ij} = 0$ for all $j \neq \ell$.

Here are three examples of matrices in reduced row echelon form

$$\mathbf{A}_{\mathrm{rref}} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ; \mathbf{B}_{\mathrm{rref}} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} ; \mathbf{C}_{\mathrm{rref}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Reaching a Generalized Reduced Row Echelon Form

Consider an $m \times n$ matrix $\mathbf{C}$
that is already in generalized row echelon form.

Suppose it has r leading non-zero elements c_{kj_k}
in rows $k = 1, 2, \ldots, r$, where j_k is increasing in k.

Starting at the pivot element $c_{rj_r} \neq 0$ in the last pivot row r,
zeroize all the elements in column j_r above this element
by subtracting from each row k above r
the multiple c_{kj_r}/c_{rj_r} of row r of the matrix $\mathbf{C}$,
while leaving row r itself unchanged.

Each of these operations of subtracting one row from a higher row
corresponds to an upper unitriangular $m \times m$ matrix,
as does the whole pivoting process.

Repeat this operation for each of the pivot elements c_{kj_k},
working from $c_{r-1,j_{r-1}}$ all the way back and up to c_{1j_1}.

The combined procedure for all the r pivot elements constructs
one upper unitriangular $m \times m$ matrix $\mathbf{U}$
such that $\mathbf{UC}$ is in GRREF.

Permuting the Columns

We have shown how to take a general $m \times n$ matrix $\mathbf{A}$ and transform it into a matrix $\mathbf{G} = \mathbf{ULPA}$ in GRREF form by applying the product of three $m \times m$ matrices:

1. an upper unitriangular matrix $\mathbf{U}$;
2. a lower unitriangular matrix $\mathbf{L}$;
3. a permutation matrix $\mathbf{P}$.

Denote its r leading non-zero elements in rows $k = 1, 2, \ldots, r$ by g_{kj_k}, where j_k is increasing in k.

We finally post multiply $\mathbf{G}$ by an $n \times n$ permutation matrix $\tilde{\mathbf{P}}$ that moves column j_k to column k, for $k = 1, 2, \ldots, r$.

It also partitions the matrix columns into two sets:

1. first, a complete set of r columns containing all the r pivots, with one pivot in each row and one in each column;
2. then second, the remaining $n - r$ columns without any pivots.

So the resulting matrix $\mathbf{G}\tilde{\mathbf{P}}$ has a diagonal sub-matrix $D_{r \times r}$ in its top left-hand corner; its diagonal elements are the pivots.

A Partly Diagonalized Matrix

Our constructions have led to the equality

$$\mathbf{G}\tilde{\mathbf{P}} = \mathbf{ULPA}\tilde{\mathbf{P}} = \begin{pmatrix} \mathbf{D}_{r \times r} & \mathbf{B}_{r \times (n-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (n-r)} \end{pmatrix}$$

The right-hand side is a partitioned $m \times n$ matrix, whose four sub-matrices have the indicated dimensions.

We may call it a "partly diagonalized" matrix.

Provided we can show that the non-negative integer $r \leq m$ is unique, independent of what pivots are chosen, we may want to call r the pivot rank of the matrix $\mathbf{A}$.

Decomposing an $m \times n$ Matrix

Premultiplying the equality

$$\mathbf{ULPA\tilde{P}} = \begin{pmatrix} \mathbf{D}_{r\times r} & \mathbf{B}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}$$

by the inverse matrix $(\mathbf{ULP})^{-1} = \mathbf{P}^{-1}\mathbf{L}^{-1}\mathbf{U}^{-1}$,
which certainly exists, gives

$$\mathbf{A\tilde{P}} = (\mathbf{ULP})^{-1} \begin{pmatrix} \mathbf{D}_{r\times r} & \mathbf{B}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}$$

Postmultiplying the result by $\mathbf{\tilde{P}}^{-1}$ leads to

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{L}^{-1}\mathbf{U}^{-1} \begin{pmatrix} \mathbf{D}_{r\times r} & \mathbf{B}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix} \mathbf{\tilde{P}}^{-1}$$

This is a decomposition of $\mathbf{A}$ into the product
of five matrices that are much easier to manipulate.

A Final Reduction

Premultiply our last partly diagonalized $m \times n$ matrix

$$\mathbf{G}\tilde{\mathbf{P}} = \mathbf{ULPA}\tilde{\mathbf{P}} = \begin{pmatrix} \mathbf{D}_{r\times r} & \mathbf{B}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}$$

by the $m \times m$ diagonal matrix

$$\mathbf{D}^* := \mathbf{diag}(d_{11}^{-1}, d_{22}^{-1}, \ldots, d_{rr}^{-1}, 0, 0, \ldots, 0)$$

whose partitioned form is $\begin{pmatrix} \mathbf{D}_{r\times r}^{-1} & \mathbf{0}_{r\times(m-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(m-r)} \end{pmatrix}$.

The result is

$$\mathbf{D}^*\mathbf{G}\tilde{\mathbf{P}} = \mathbf{D}^*\mathbf{ULPA}\tilde{\mathbf{P}} = \begin{pmatrix} \mathbf{I}_r & \mathbf{B}^*_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}$$

where $\mathbf{B}^*_{r\times(n-r)} := (\mathbf{D}_{r\times r})^{-1}\mathbf{B}_{r\times(n-r)}$.

So the diagonal matrix in the top left corner
has been converted to the identity.

Special Cases

So far we have been writing out full partitioned matrices, as is required when the number of pivots satisfies $r < \min\{m, n\}$.

Here are three other special cases when $r \geq \min\{m, n\}$, where the partially diagonalized $m \times n$ matrix

$$\mathbf{G}\tilde{\mathbf{P}} = \mathbf{ULPA}\tilde{\mathbf{P}} = \begin{pmatrix} \mathbf{D}_{r\times r} & \mathbf{B}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}$$

reduces to:

1. $\begin{pmatrix} \mathbf{D}_{m\times m} & \mathbf{B}_{m\times(n-m)} \end{pmatrix}$ in case $r = m < n$, so $m - r = 0$;
2. $\begin{pmatrix} \mathbf{D}_{n\times n} \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix}$ in case $r = n < m$, so $n - r = 0$;
3. $\mathbf{D}_{n\times n}$ in case $r = m = n$, so $m - r = n - r = 0$.

Outline

Finding the Determinant of a Square Matrix

In the case of an $n \times n$ matrix $\mathbf{A}$, our earlier equality becomes

$$\mathbf{ULPA\tilde{P}} = \begin{pmatrix} \mathbf{D}_{r\times r} & \mathbf{B}_{r\times(n-r)} \\ \mathbf{0}_{(n-r)\times r} & \mathbf{0}_{(n-r)\times(n-r)} \end{pmatrix}$$

The determinant of this upper triangular matrix is clearly 0 except in the special case when $r = n$.

When $r = n$, there is a complete set of n pivots.

There are no missing columns, so no need to permute the columns by applying the permutation matrix $\tilde{\mathbf{P}}$.

Instead, we have the complete diagonalization $\mathbf{ULPA} = \mathbf{D}$.

The unitriangular matrices have determinants $|\mathbf{U}| = |\mathbf{L}| = 1$.

Also $|\mathbf{P}| = |\mathbf{P}^{-1}| = \pm 1$, depending on the common sign of the permutation $\mathbf{P}$ and its inverse.

So the product rule for determinants implies that $|\mathbf{A}| = \text{sgn}(\mathbf{P})|\mathbf{D}|$.

It is enough to multiply the diagonal elements, and choose the sign.

A Matrix Equation

Consider the matrix equation $\mathbf{AX} = \mathbf{Y}$ where

1. the matrix $\mathbf{A}$ is $m \times n$;
2. the matrix $\mathbf{X}$ is $n \times p$;
3. the matrix $\mathbf{Y}$ is $m \times p$.

Really, it is p systems of m equations in n unknowns.

Premultiplying by $\mathbf{D}^*\mathbf{ULP}$, then manipulating, transforms the left-hand side of the matrix equation $\mathbf{AX} = \mathbf{Y}$ to

$$\mathbf{D}^*\mathbf{ULPAX} = \mathbf{D}^*\mathbf{ULPA}\tilde{\mathbf{P}}\tilde{\mathbf{P}}^{-1}\mathbf{X} = \begin{pmatrix} \mathbf{I}_{r\times r} & \mathbf{B}^*_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix} \tilde{\mathbf{P}}^{-1}\mathbf{X}$$

So the whole equation $\mathbf{AX} = \mathbf{Y}$ gets transformed to

$$\begin{pmatrix} \mathbf{I}_{r\times r} & \mathbf{B}^*_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix} \tilde{\mathbf{P}}^{-1}\mathbf{X} = \mathbf{D}^*\mathbf{ULPY}$$

Inverting a Square Matrix

Suppose that $\mathbf{A}$ is $n \times n$,
and consider the equation system $\mathbf{AX} = \mathbf{I}_n$.

It has a solution if and only if $|\mathbf{A}| \neq 0$,
in which case there is a unique solution $\mathbf{X} = \mathbf{A}^{-1}$.

The necessary and sufficient condition $|\mathbf{A}| \neq 0$ for invertibility holds if and only if there is a full set of n pivots, so $\mathbf{ULPA} = \mathbf{D}$.

Then $\mathbf{AX} = \mathbf{I}_n$ implies that $\mathbf{ULPAX} = \mathbf{DX} = \mathbf{ULPI} = \mathbf{ULP}$.

So $\mathbf{X} = \mathbf{A}^{-1} = \mathbf{D}^{-1}\mathbf{ULP}$.

Pivoting does virtually all the work of matrix inversion,
because all that is left to invert a diagonal matrix,
then find the product of four $n \times n$ matrices.

Of these four matrices, one is diagonal, two are triangular,
and one is a permutation.