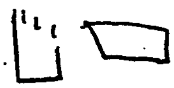
GAUSS FACTORIZATION
TABLE OF CONTENTS

GAUSS FACTORIZATION

LU Theorem.

Let $O \neq A \in \mathbb{F}^{n \times m}$. There are permutation matrices P and Q, and an integer ρ with $1 \leq \rho \leq \min\{m, n\}$, so that

$$Q'AP = LU =$$ 

with $L \in \mathbb{F}^{n \times \rho}$ *unit* lower trapezoidal and $U \in \mathbb{F}^{\rho \times m}$ upper trapezoidal with nonzero diagonal elements. Furthermore, P and Q can be chosen so that all elements of L are at most 1 in absolute value.

□ By induction on (actually, reduction of) min {m, n}.

a. Since $A \neq O$ there are interchange (i.e., special permutation) matrices P_1 and Q_1 so that

$$Q_1'AP_1 = \begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix}, \quad \alpha \neq 0 .$$

α is the first *pivot*. Factor

$$\begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix} = \begin{bmatrix} 1 & \\ \ell & I \end{bmatrix} \begin{bmatrix} \alpha & b' \\ & G \end{bmatrix}.$$

This requires

$$a = \ell\alpha, \qquad B = \ell b' + G .$$

Thus we compute

$$\ell = \frac{a}{\alpha}, \qquad G = B - \ell b' .$$

G is called the *Gauss transform* of $\begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix}$. It is also called, more commonly, the *Schur complement* of the pivot α in $\begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix}$.

We then have

$$Q_1'AP_1 = \begin{bmatrix} 1 & \\ \ell & I \end{bmatrix} \begin{bmatrix} \alpha & b' \\ & G \end{bmatrix}$$

with $G \in \mathbb{F}^{(n-1)\times(m-1)}$.

The *complete pivoting* strategy chooses P_1 and Q_1 so that $|\alpha| = \max |\alpha_{ji}|$. This guarantees that all elements of $\ell = \frac{a}{\alpha}$ are ≤ 1 in absolute value.

b. If $G = O$, or if $G = \emptyset$ is the empty matrix, we put

$$P = P_1, \qquad Q = Q_1, \qquad \rho = 1,$$

and *stop*, with

$$Q'AP = \begin{bmatrix} 1 \\ \ell \end{bmatrix} \begin{bmatrix} \alpha & b' \end{bmatrix} = LU,$$

as required. Note that $G = \emptyset \Leftrightarrow B = \emptyset$. This can happen in three ways:

i. $m = 1 \Leftrightarrow b' = \emptyset \Rightarrow$

$$Q'AP = \begin{bmatrix} 1 \\ \ell \end{bmatrix} \begin{bmatrix} \alpha \end{bmatrix} = LU,$$

ii. $n = 1 \Leftrightarrow a = \emptyset \Rightarrow$

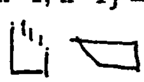
$$Q'AP = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} \alpha & b' \end{bmatrix} = LU;$$

iii. $m = n = 1 \Leftrightarrow a = \emptyset,\ b = \emptyset \Rightarrow$

$$Q'AP = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} \alpha \end{bmatrix} = LU.$$

This shows that the induction assumption holds when $\min\{m, n\} = 1$.

c. Now $O \neq G \in \mathbb{F}^{(n-1)\times(m-1)}$. By the induction hypothesis there are permutation matrices P_2 and Q_2, and an integer σ with $1 \leq \sigma \leq \min\{m-1, n-1\} = \min\{m, n\} - 1$, so that

$$Q_2'GP_2 = MV =$$ 

with $M \in \mathbb{F}^{(n-1)\times\sigma}$ unit lower trapezoidal and $V \in \mathbb{F}^{\sigma\times(m-1)}$ upper trapezoidal with nonzero diagonal elements. Moreover, we can choose P_2 and Q_2 so all elements of M are ≤ 1 in absolute value.

d. We then have

$$\begin{bmatrix} 1 & \\ & Q_2' \end{bmatrix} Q_1'AP_1 \begin{bmatrix} 1 & \\ & P_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ & Q_2' \end{bmatrix} \begin{bmatrix} 1 & \\ \ell & I \end{bmatrix} \begin{bmatrix} \alpha & b' \\ & G \end{bmatrix} \begin{bmatrix} 1 & \\ & P_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ Q_2'\ell & Q_2' \end{bmatrix} \begin{bmatrix} \alpha & b'P_2 \\ & GP_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ Q_2'\ell & I \end{bmatrix} \begin{bmatrix} 1 & \\ & Q_2' \end{bmatrix} \begin{bmatrix} \alpha & b'P_2 \\ & GP_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ Q_2'\ell & I \end{bmatrix} \begin{bmatrix} \alpha & b'P_2 \\ & Q_2'GP_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ Q_2'\ell & I \end{bmatrix} \begin{bmatrix} \alpha & b'P_2 \\ & MV \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ Q_2'\ell & I \end{bmatrix} \begin{bmatrix} 1 & \\ & M \end{bmatrix} \begin{bmatrix} \alpha & b'P_2 \\ & V \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ Q_2'\ell & M \end{bmatrix} \begin{bmatrix} \alpha & b'P_2 \\ & U \end{bmatrix} =: LU ,$$

as required, since

$$P := P_1 \begin{bmatrix} 1 & \\ & P_2 \end{bmatrix}, \quad Q := Q_1 \begin{bmatrix} 1 & \\ & Q_2 \end{bmatrix}$$

are permutation matrices. We have $\rho = 1 + \sigma \le \min\{m, n\}$. ■

We shall *soon* show that

$$\rho = \rho(A),$$

the number of pivots, is independent of the pivot strategy!

Our matlab code gfr *implements* the above proof. We must include a tolerance, tol, to cope with rounding errors.

function [L, U, p, q, g] = gfr (A, tol)

Gauss factorization with complete pivoting for size and *recursive* calls.

Illustrates the *reductive* nature of the Gauss factorization process and provides an implementation of the most fundamental proof of the LU theorem.

Displays intermediate output. Press any key to continue.

 Revised 20 April 1994.

gfr calls cols and gfr.

begin gfr

```
[n  m] = size(A);      r = min(m, n);
if r < 1
   g = 1;     return
end
```

Now A is not the empty matrix. Find the pivot.

```
[a  t] = max(abs(A));     [a  s] = max(a);     t = t(s);
if n < 2
   s = t;     t = 1;
end
```

Do special things if this is the first call to gfr.

```
if nargin < 2
   tol = r*a;     format compact,     format short,     disp(' ')
end
A,     pause
```

The tolerance tol is computed at the highest level and passed as an argument to the lower levels. If a is negligible with respect to tol we are done. Then L and U are the empty matrix.

```
If tol + a == tol
    p = 1:m;     q = 1:n;     g = a;
else
```

Now A ~= 0. "Do Gauss". It is not necessary to physically make the interchanges. It is also more convenient, here, to reorder columns 1, 2, ..., s as s, 1, 2, ..., s−1 and, likewise, to reorder rows 1, 2, ..., t as t, 1, 2, ..., t−1. In extreme cases, when there are ties for the pivots, this can lead to LU factorizations which are different from those of gf.

```
        i = [1:s-1  s+1:m];                                  j = [1:t-1  t+1:n]
        u = A(t, i)';             pivot = A(t, s);           ℓ = A(j, s)/pivot;
        p = [s  i];                                          q = [t  j];
        if min (m, n)  <  2
            This is the case when gfr need not be called again.
            L = [1;  ℓ];     U = [pivot  u'];     g = a:
        else
            Compute the (nonempty) Gauss transform, again called A, with one less column
            and row than the input A.  Call gfr again, and in fact recursively, with this A.
            Use the results to build the output for the current call to gfr.  The unscaled
            growth factor g is passed from the lower level calls to gfr to the higher ones.
            It is scaled only at the last step, that is only in the initial call to gfr.
            A = A(j, i) - l*u';      [L  U  s  t  g] = gfr(A, tol);     g = max(g, a);
            r = cols(L);             z = zeros(r, 1);
            L = [1  z';  ℓ(t)  L];   U = [pivot u(s)';  z  U];
            p = p([1  s+1]);         q = q([1  t+1]);
        end
        L,     U,     pause
    end
    Scale g by a if this is the highest level call to gfr.
    if nargin < 2
      if a  > 0
        g = g/a;
      else
        g = 1;
      end
    end
and gfr
```

Example

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix}, \qquad m = 3, n = 2\,.$$

a. *No pivoting.* We now view the array containing A as a "tableau," or "workspace," and not as a matrix. Pivots are circled.

$$\left[\begin{array}{c|cc} 1 & 0 & 1 \\ \hline 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 1 & -3 & -3 \\ \hline 1 & 1 & 1 \\ 1 & -2 & -2 \end{array}\right]$$

We

- copy the pivot and what's to its right (b'),
- divide what's below the pivot by the pivot, thus forming $\ell = \frac{a}{\alpha}$,
- replace B by $B - \ell b'$, circling the next pivot.

We then repeat the process, as long as we can find a pivot. By definition, *pivots are not zero.*

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 1 & -3 & -3 \\ \hline 1 & 1 & 1 \\ 1 & -2 & -2 \end{array}\right] \to \left[\begin{array}{ccc} 1 & 0 & 1 \\ 1 & -3 & -3 \\ 1 & -1/3 & 0 \\ 1 & 2/3 & 0 \end{array}\right]$$

Here, there are no more pivots. The nontrivial part of L, excluding the "automatic" 1s and 0s, is beneath the "staircase." The nontrivial part of U is on the "staircase" and above. *Result:* $\rho = 2$ and

$$A = LU = \begin{bmatrix} 1 & \\ 1 & 1 \\ 1 & -1/3 \\ 1 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ & -3 & -3 \end{bmatrix}.$$

All such results should be *checked.* That's easy here.

b. *Complete pivoting.* We now use an *extended* tableau which now also contains the column and row indices, which will ultimately be the permutations p and q. This leaves one extra space which we use as a "stage counter" and which will ultimately contain ρ. We view the interchange steps as $\frac{1}{2}$-stages. We interchange to put the pivot in the pivot position. We choose the pivot to be the *first* element of largest absolute value in the columnwise ordering (A(:) in matlab). Note that rounding errors can cause mathematical ties to be resolved differently than with pen(cil) and paper computations.

If the matrix is square ($n \times n$) and there are n pivots, there is still an nth stage, even though there is nothing to the right of, or below, the pivot. This last stage consists merely of recognizing that there *is* a pivot , and incrementing the stage counter.

Here are the computations for the given A:

$$
\begin{array}{c|ccc}
0 & 1 & 2 & 3 \\ \hline
1 & 1 & 0 & 1 \\
2 & 1 & -3 & -2 \\
3 & 1 & 1 & 2 \\
4 & 1 & -2 & -1
\end{array}
\quad \longrightarrow \quad
\begin{array}{c|c:cc}
1/2 & 2 & 1 & 3 \\ \hline
2 & -3 & 1 & -2 \\ \hline
1 & 0 & 1 & 1 \\
3 & 1 & 1 & 2 \\
4 & -2 & 1 & -1
\end{array}
\quad \longrightarrow
$$

$$
\begin{array}{c|c|cc}
1 & 2 & 1 & 3 \\ \hline
2 & -3 & 1 & -2 \\ \hline
1 & 0 & 1 & 1 \\
3 & -1/3 & 4/3 & 4/3 \\
4 & 2/3 & 1/3 & 1/3
\end{array}
\quad \longrightarrow \quad
\begin{array}{c|cc|c}
3/2 & 2 & 1 & 3 \\ \hline
2 & -3 & 1 & -2 \\
3 & -1/3 & 4/3 & 4/3 \\ \hline
1 & 0 & 1 & 1 \\
4 & 2/3 & 1/3 & 1/3
\end{array}
\quad \longrightarrow
$$

$$
\begin{array}{c|cc|c}
2 & 2 & 1 & 3 \\ \hline
2 & -3 & 1 & -2 \\
3 & -1/3 & 4/3 & 4/3 \\ \hline
1 & 0 & 3/4 & 0 \\
4 & 2/3 & 1/4 & 0
\end{array}
\qquad \rho = 2, \qquad \mathrm{p} = [2\ \ 1\ \ 3], \qquad \mathrm{q} = [2\ \ 3\ \ 1\ \ 4],
$$

Results. ρ, p and q are given just above, and we have

$$
\mathrm{Q'AP} = \mathrm{LU} = \begin{bmatrix} 1 & \\ -1/3 & 1 \\ 0 & 3/4 \\ 2/3 & 1/4 \end{bmatrix} \begin{bmatrix} -3 & 1 & -2 \\ & 4/3 & 4/3 \end{bmatrix}
$$

Check:

$$
\mathrm{LU} = \begin{bmatrix} -3 & 1 & -2 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \\ -2 & 1 & -1 \end{bmatrix},
$$

$$
Q'AP = Q' \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} e_2 & e_1 & e_3 \end{bmatrix}
$$

$$
= \begin{bmatrix} e_2' \\ e_3' \\ e_1' \\ e_4' \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ -3 & 1 & -2 \\ 1 & 1 & 2 \\ -2 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -3 & 1 & -2 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \\ -2 & 1 & -1 \end{bmatrix} = LU.
$$

The codes which implement this ***nonrecursive*** form of the algorithm are gfpc, a demonstration code, and gf, our "production" code. You can check all your results with gfpc, if $m \le 7$ and there are no ties for the pivots. Also, gfpn does ***no pivoting*** and gfps chooses ***small pivots***, just to show the "numerical instability" of this scheme.

"*Solving*" $Ax = b$.

We have the following equivalences:

$$
\begin{aligned}
Ax = b \quad &\Leftrightarrow \quad Q'Ax = Q'b && \text{since } QQ' = I, \\
&\Leftrightarrow \quad Q'AP \cdot P'x = Q'b && \text{since } PP' = I, \\
&\Leftrightarrow \quad LU \cdot P'(x) = Q'b && \text{since } Q'AP = LU, \\
&\Leftrightarrow \quad Lc = Q'b, \quad U \cdot P'x = c, && \text{definition of } c.
\end{aligned}
$$

Partition

$$
LU = \begin{array}{c} \rho \\ n-\rho \end{array} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix} \rho
$$

$$
\qquad \quad \rho \qquad \quad \rho \quad m-\rho
$$

and

$$
b(q) := Q'b =: \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{array}{c} \rho \\ n-\rho \end{array},
$$

$$
x(p) := P'x =: \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{array}{c} \rho \\ m-\rho \end{array}.
$$

Thus we have

$$\boxed{Ax = b \quad \Leftrightarrow \quad L_1 c = b_1, \qquad L_2 c = b_2, \qquad U_1 x_1 + U_2 x_2 = c}\ .$$

$L_1 =$ is unit lower triangular. $U_1 =$ is upper triangular with nonzero diagonal elements (the pivots).

Either of the matrices L_2 or U_2, or both, and the corresponding vectors b_2 or x_2, or both, may be empty ($\emptyset$). We have

$$\rho = m \quad \Leftrightarrow \quad U_2 = \emptyset, \qquad x_2 = \emptyset, \qquad P'x = x_1,$$

and

$$\rho = n \quad \Leftrightarrow \quad L_2 = \emptyset, \qquad b_2 = \emptyset, \qquad Q'b = b_1\ .$$

We can now "solve" $Ax = b$ completely by using only triangular solves involving L_1 and U_1, and matrix-vector multiplications involving L_2 and U_2, if they are not empty.

Since L_1 has nonzero diagonal elements we can forward solve the system $L_1 c = b_1$, uniquely, for c. The linear system $Ax = b$ is then *solvable* if and only if this c satisfies the *solvability condition*

$$\boxed{L_2 c = b_2}\ .$$

If $\rho = n$ this condition does not arise. Thus, if $\rho = n$ then $Ax = b$ is solvable, for *every* $b \in \mathbb{R}^n$. If $\rho < n$ and the solvability condition is *not* satisfied then we *stop*: $Ax = b$ *has no solution.*

We must now solve

$$U_1 x_1 + U_2 x_2 = c\ .$$

If $\rho = m$ this is just

$$U_1 x_1 = c$$

which we backsolve, uniquely, for x_1 and then put $x = Px_1$. Thus, if $\rho = m$ *the* solution of $Ax = b$, if one exists, is unique.

If $\rho < m$ then we may assign *arbitrary* values to the elements of $x_2 \in \mathbb{R}^{m-\rho}$. These elements are usually called "free variables." There are $m-\rho$ of them. If $\rho = m$ there are none of them! Once the free variables are assigned we backsolve

$$U_1 x_1 = c - U_2 x_2$$

uniquely for x_1. Then we "unshuffle" to get

$$x = P\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

If the free variables are retained as *variables* then this x is the *general solution* of $Ax = b$. *The free variables may be different if different pivot strategies are used.*

If the free variables are given specific values, then we obtain a *particular solution* of $Ax = b$. The simplest particular solution is obtained by taking $x_2 = 0_{m-\rho}$. This is the particular solution which is computed by our code gfsp. We b-solve $Ux_1 = c$ for x_1 and then put $x = P\begin{bmatrix} x_1 \\ 0 \end{bmatrix}$. In matlab one does this by writing x(p) = [x1; 0]. If $Ax = b$ is not solvable we output x := [ ]; matlab's empty vector (matrix). The solvability question $L_2c \overset{?}{=} b_2$ is normally obscured by rounding errors. We use a "reasonable" criterion for replacing "tiny" $b_2 - L_2c$ by $o_{n-\rho}$, if $\rho < n$.

We shall give examples soon. But first, while the *general* discussion is at hand, we need to establish some *basic facts.*

First of all, the *homogeneous* system

$$Ax = 0_n$$

is always solvable, by $x = 0_m$. This is the unique solution of $Ax = 0_n$ if and only if $\rho = m$ (no free variables). Otherwise there are free variables and we can choose any, or all, of them to be nonzero.

Basic fact GF1. The homogeneous system $Ax = 0_n$ has a nonzero solution if and only if (there is a pivot strategy with) $\rho < m$.

The strongest consequence of this which is independent of ρ is

Key fact GF2. Let $A \in \mathbb{F}^{n \times m}$. If $m > n$ then the homogeneous system $Ax = 0_n$ has a nonzero solution.

□ Always $\rho \le \min\{m, n\}$. If $m > n$ then $\rho \le \min\{m, n\} = n < m$. Now apply fact GF1. ■

It is conceivable that, for a given matrix A, different pivot strategies could give rise to different numbers of pivots, ρ. For instance, for $A \in \mathbb{F}^{n \times n}$ there could be $(n!)^2$ different ways to choose the pivots.

Basic fact GF3. $\rho = \rho(A)$ depends only on A, and not on the pivot strategy used to factor A.

□ Consider the homogeneous system $Ax = 0_n$. With one pivot strategy this is *equivalent with*

$$UP'x = 0, \qquad U \in \mathbb{R}^{\rho \times m}$$

and, with another, with

$$\mathrm{V}\widehat{\mathrm{P}}'\mathrm{x} = 0, \qquad \mathrm{V} \in \mathbf{F}^{\sigma \times m}.$$

The matrices U and V are upper trapezoidal with nonzero diagonal elements. We wish to prove that $\rho = \sigma$.

If not, we may assume that $\rho > \sigma$. Now the two systems $\mathrm{UP}'\mathrm{x} = \underset{\sim}{0}$ and $\mathrm{V}\widehat{\mathrm{P}}'\mathrm{x} = 0$ have the *same* solution sets, namely the solution set $\mathcal{S}$ of $\mathrm{Ax} = 0_{\mathrm{n}}$.

Suppose first that $\rho = \mathrm{m}$. Then the U-system has the unique solution $\mathrm{x} = 0_{\mathrm{m}}$ so $\mathcal{S} = \{0_{\mathrm{n}}\}$. On the other hand the V-system is a homogeneous linear system of $\sigma < \rho = \mathrm{m}$ equations in m unknowns. By key fact GF2 it has a nonzero solution so $\mathcal{S} \neq \{0_{\mathrm{m}}\}$. This contradiction proves our result when $\rho = \mathrm{m}$.

Now suppose $\rho < \mathrm{m}$. Then the U-system is

$$\mathrm{U}_1\mathrm{x}_1 + \mathrm{U}_2\mathrm{x}_2 = 0, \qquad \mathrm{x} = \mathrm{P}\begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \end{bmatrix},$$

with $\mathrm{x}_2 \in \mathbb{R}^{\mathrm{m}-\rho}$. *Append* to this system the $\mathrm{m}-\rho$ equations

$$\mathrm{x}_2 = 0_{\mathrm{m}-\rho}\,.$$

The solution set of this appended U-system is $\mathcal{S} = \{0_{\mathrm{m}}\}$, and it is the *same* set as the solution set of the *appended* V-system:

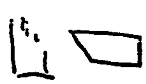
$$\mathrm{V}\widehat{\mathrm{P}}'\mathrm{x} = 0, \qquad \mathrm{x}_2 = 0_{\mathrm{m}-\rho}\,.$$

But this is a homogeneous system of $\sigma + \mathrm{m} - \rho = \mathrm{m} - (\rho - \sigma) < \mathrm{m}$ equations in m unknowns, so $\mathcal{S} \neq \{0_{\mathrm{m}}\}$ by key fact GF2. This contradiction proves our result also when $\rho < \mathrm{m}$. ■

The parenthetical expression in basic fact GF1 may now be deleted.

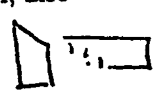
The LU factorization is a "*dual*" factorization. That is if

$$\mathrm{Q}'\mathrm{AP} = \mathrm{LU} =$$



then, by the reverse order rule for transposition, also

$$\mathrm{P}'\mathrm{A}'\mathrm{Q} = \mathrm{U}'\mathrm{L}' =$$



The placement of the 1s on the diagonal of L is not important. We could as well have made the diagonal elements of U 1s. (How may this be done, after the fact?) What *is* important is that the diagonal elements of L *and* U are *not zero*. The above shows that if A has such a factorization then so does A′, and vice versa. In other words, with "Gauss," whatever we do for A we can also do A′, "dually." In particular we have

Basic fact GF4. $\rho(A') = \rho(A)$.

We now recall our setup and give some examples. With the partitionings

$$Q'AP = LU = \begin{array}{c} \rho \\ n-\rho \\ \end{array} \begin{array}{c} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \\ \rho \end{array} \begin{array}{cc} \begin{bmatrix} U_1 & U_2 \end{bmatrix} & \rho \\ \rho \quad m-\rho & \end{array}$$

and

$$Q'b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{array}{c} \rho \\ n-\rho \end{array}$$

$$P'x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{array}{c} \rho \\ m-\rho \end{array}$$

we execute

Algorithm "solve".

1. f-solve $L_1c = b_1$ for c.

2. Check *solvability:*

$$L_2c \stackrel{?}{=} b_2 \, .$$

   If not solvable, stop.

3. If $\rho = m$ the *unique* solution is

$$x = Px_1: \qquad Ux_1 = c \, .$$

   If $\rho < m$ the *general* solution is

$$x = P\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}: \qquad U_1x_1 = c - U_2x_2$$

   with $x_2 \in \mathbb{F}^{m-\rho}$ *arbitrary.*

Examples. On pages 5-8 we factored

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix}$$

in two ways, with no pivoting and with complete pivoting. We now consider two different right sides

$$b = Ax_0 = \begin{bmatrix} 4 \\ -11 \\ 9 \\ -6 \end{bmatrix}: \quad x_0 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

and

$$b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

We execute algorithm "solve" for the two different LU factorizations and these two bs. We find the *general* solution to $Ax = b$.

1. *No pivoting.* $\rho = 2$,

$$A = LU = \left[\begin{array}{cc} 1 & \\ 1 & 1 \\ \hline 1 & -1/3 \\ 1 & 2/3 \end{array}\right] \left[\begin{array}{cc|c} 1 & 0 & 1 \\ & -3 & -3 \end{array}\right] = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix}.$$

a. $$b = \left[\begin{array}{c} 4 \\ -11 \\ \hline 9 \\ -6 \end{array}\right] = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

$$L_1 c = b_1: \quad \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 4 \\ -11 \end{bmatrix}, \quad c = \begin{bmatrix} 4 \\ -15 \end{bmatrix}$$

$L_2c \overset{?}{=} b_2$: $\quad L_2c = \begin{bmatrix} 1 & -1/3 \\ 1 & -2/3 \end{bmatrix} \begin{bmatrix} 4 \\ -15 \end{bmatrix} = \begin{bmatrix} 9 \\ 6 \end{bmatrix} = b_2 .$

Solvable. Proceed.

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \hline \xi_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$U_1x_1 = c - U_2x_2$:

$$\begin{bmatrix} 1 & 0 \\ & -3 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 4 \\ -15 \end{bmatrix} - \begin{bmatrix} 1 \\ -3 \end{bmatrix} \xi_3$$

$$= \begin{bmatrix} 4 - 1\xi_3 \\ -15 + 3\xi_3 \end{bmatrix}.$$

b-solve for ξ_1, ξ_2, as functions of the free variable ξ_3:

$$-3\xi_2 = -15 + 3\xi_3, \qquad \xi_2 = 5 - \xi_3, \qquad \xi_1 = 4 - 1\xi_3 .$$

Thus the *general solution* is

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} 4 - 1\xi_3 \\ 5 - 1\xi_3 \\ 0 + 1\xi_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \xi_3$$

$$= x_p + Mz \qquad (z = \xi \in \mathbb{R} \text{ arbitrary})$$

The particular solution given by gfsp is x_p ($\xi_3 = 0$). This is *not* the solution x_0 we built b from! But there is a unique value of ξ_3 for which $x = x_0$, namely $\xi_3 = 3$. Geometrically, the general solution is a line in $\mathbb{R}^3$ passing through x_p and x_0. Otherwise stated, the general solution is all scalar multiples of the nonzero vector M (a line through 0_3) *translated* to pass through x_p, or x_0. *Any nonzero* scalar multiple of M would serve as well to determine this line, since the *free variable* ξ_3 varies over all of $\mathbb{R}$.

Check:

$$Ax_p = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ -11 \\ 9 \\ -6 \end{bmatrix} = b\,,$$

$$AM = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0_4\,.$$

Thus

$$\begin{aligned} Ax &= A(x_p + Mz) \\ &= Ax_p + AMz \\ &= b + 0z \\ &= b\,. \end{aligned}$$

The set $\{Mz\colon\ z = \xi \in \mathbf{R}\}$ is a line through 0_3. It consists of all solutions of the homogeneous system $Ax = 0_n$.

b.

$$b = \left[\begin{array}{c} 1 \\ 2 \\ \hline 3 \\ 4 \end{array}\right] = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

$L_1c = b_1$:

$$\begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

$L_2c \stackrel{?}{=} b_2$:

$$L_2c = \begin{bmatrix} 1 & -1/3 \\ 1 & -2/3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} \neq b_2\,.$$

Not solvable. *Stop!*

2. *Complete pivoting.* Also $\rho = 2$,

$$Q'AP = LU = \left[\begin{array}{cc} 1 & \\ -1/3 & 1 \\ \hline 0 & 3/4 \\ 2/3 & 1/4 \end{array}\right] \left[\begin{array}{cc|c} -3 & 1 & -2 \\ & 4/3 & 4/3 \end{array}\right] = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix},$$

with $p = \begin{bmatrix} 2 & 1 & 3 \end{bmatrix}$ and $q = \begin{bmatrix} 2 & 3 & 1 & 4 \end{bmatrix}$.

a.
$$b = \begin{bmatrix} 4 \\ -11 \\ 9 \\ -6 \end{bmatrix}, \quad b(q) = Q'b = \left[\begin{array}{c} -11 \\ 9 \\ \hline 4 \\ -6 \end{array}\right] = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$L_1c = b_1$:
$$\begin{bmatrix} 1 & \\ -1/3 & 1 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} -11 \\ 9 \end{bmatrix}, \quad c = \begin{bmatrix} -11 \\ 16/3 \end{bmatrix}.$$

$L_2c \overset{?}{=} b_2$:
$$L_2c = \begin{bmatrix} 0 & 3/4 \\ 2/3 & 1/4 \end{bmatrix} \begin{bmatrix} -11 \\ 16/3 \end{bmatrix} = \begin{bmatrix} 4 \\ -6 \end{bmatrix} = b_2.$$

Solvable. Proceed.

$$x(p) = P'x = \left[\begin{array}{c} \xi_2 \\ \xi_1 \\ \hline \xi_3 \end{array}\right] = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$U_1x_1 = c - U_2x_2$:

$$\begin{bmatrix} -3 & 1 \\ & 4/3 \end{bmatrix} \begin{bmatrix} \xi_2 \\ \xi_1 \end{bmatrix} = \begin{bmatrix} -11 \\ 16/3 \end{bmatrix} - \begin{bmatrix} -2 \\ 4/3 \end{bmatrix} \xi_3$$

$$= \begin{bmatrix} -11 + 2\xi_3 \\ 16/3 - 4/3\ \xi_3 \end{bmatrix}.$$

b-solve for ξ_1, ξ_2 as functions of the free variable ξ_3:

$$\tfrac{4}{3}\xi_1 = \tfrac{16}{3} - \tfrac{4}{3}\xi_3, \qquad \xi_1 = 4 - 1\xi_3$$

$$-3\xi_2 + \xi_1 = -11 + 2\xi_3 \,,$$

$$-3\xi_2 = -\xi_1 - 11 + 2\xi_3 \,,$$

$$= -4 = \xi_3 - 11 + 2\xi_3 \,,$$

$$= -15 + 3\xi_3 \,,$$

$$\xi_2 = 5 - 1\xi_3 \,.$$

The general solution is

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} 4-\xi_3 \\ 5-\xi_3 \\ 0+\xi_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} \div \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \xi_3$$

$$= x_p + Mz \,,$$

as before. This was an accident! *Usually*, the particular solutions x_p, the "basis matrices" M, and even the vector $z := x_2$ of free variables will be different for different pivot strategies.

We illustrate this by solving the "dual" homogeneous problem

$$A'y = 0_m$$

for its general solution using the "dual" LU factorizations

$$P'A'Q' = U'L' = \begin{bmatrix} U_1' \\ U_2' \end{bmatrix} \begin{bmatrix} L_1' & L_2' \end{bmatrix}$$

of the above two factorizations. When working with A' we make the exchanges

$$A \leftrightarrow A', \qquad m \leftrightarrow n$$

$$P \leftrightarrow Q, \qquad L \leftrightarrow U' \,.$$

Of course ρ is left unchanged since

$$\rho(A) = \rho(A') \,.$$

Since $c = 0$ we have only to solve

$$L_1'y_1 = 0 - L_2'y_2$$

in each case, where $w := y_2$ is the vector of free variables and

$$y(q) = Q'y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \begin{matrix} 0 \\ n-0 \end{matrix} \,.$$

3. *No pivoting.*

$$\mathrm{y(q)} = \mathrm{Q'y} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \hline \eta_3 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} \mathrm{y}_1 \\ \mathrm{y}_2 \end{bmatrix}$$

$\mathrm{L}_1'\mathrm{y}_1 = -\mathrm{L}_2'\mathrm{y}_2$:

$$\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = -\begin{bmatrix} 1 & 1 \\ -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} \eta_3 \\ \eta_4 \end{bmatrix}$$

$$= \begin{bmatrix} -\eta_3 - \eta_4 \\ \frac{1}{3}\eta_3 - \frac{2}{3}\eta_4 \end{bmatrix}.$$

b-solve for η_1, η_2 as functions of the free variables η_3 and η_4:

$$\begin{aligned} \eta_2 &= \frac{1}{3}\eta_3 - \frac{2}{3}\eta_4 , \\ \eta_1 + \eta_2 &= -\eta_3 - \eta_4 , \\ \eta_1 &= -\eta_2 - \eta_3 - \eta_4 , \\ &= -\frac{1}{3}\eta_3 + \frac{2}{3}\eta_4 - \eta_3 - \eta_4 . \\ &= -\frac{4}{3}\eta_3 - \frac{1}{3}\eta_4 . \end{aligned}$$

Thus the *general* solution of $\mathrm{A'y} = 0$ is

$$\mathrm{y} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} -\frac{4}{3}\eta_3 - \frac{1}{3}\eta_4 \\ \frac{1}{3}\eta_3 - \frac{2}{3}\eta_4 \\ 1\eta_3 + 0\eta_4 \\ 0\eta_3 + 1\eta_4 \end{bmatrix}$$

$$= \begin{bmatrix} -4/3 & -1/3 \\ 1/3 & -2/3 \\ \hline 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_3 \\ \eta_4 \end{bmatrix} =: \mathrm{Nw}$$

Check. We have

$$A'N = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -3 & 1 & -2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \begin{bmatrix} -4/3 & -1/3 \\ 1/3 & -2/3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = O$$

so, of course,

$$A'y = A'Nw = Ow = 0.$$

4. *Complete pivoting.*

$$y(q) = Q'y = \begin{bmatrix} \eta_2 \\ \eta_3 \\ \hline \eta_1 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$L_1'y_1 = -L_2'y_2$:

$$\begin{bmatrix} 1 & -1/3 \\ & 1 \end{bmatrix} \begin{bmatrix} \eta_2 \\ \eta_3 \end{bmatrix} = -\begin{bmatrix} 0 & 2/3 \\ 3/4 & 1/4 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_4 \end{bmatrix}$$

$$= \begin{bmatrix} 0\,\eta_1 - \frac{2}{3}\eta_4 \\ -\frac{3}{4}\eta_1 - \frac{1}{4}\eta_4 \end{bmatrix}.$$

b-solve for η_2, η_3 as functions of the free variables η_1 and η_4:

$$\eta_2 = -\frac{3}{4}\,\eta_1 - \frac{1}{4}\,\eta_4\,,$$

$$\eta_2 - \frac{1}{3}\eta_3 = 0\eta_1 - \frac{2}{3}\eta_4\,,$$

$$\eta_2 = \frac{1}{3}\eta_3 - \frac{2}{3}\eta_4$$

$$= -\frac{1}{4}\eta_1 - \frac{1}{12}\eta_4 - \frac{2}{3}\eta_4$$

$$= -\frac{1}{4}\eta_1 - \frac{3}{4}\eta_4\,.$$

Thus the *general* solution of $A'y = 0$ is *also*

$$y = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} 1\eta_1 + 0\eta_4 \\ -\frac{1}{4}\eta_1 - \frac{3}{4}\eta_4 \\ -\frac{3}{4}\eta_1 - \frac{1}{4}\eta_4 \\ 0\eta_1 + 1\eta_4 \end{bmatrix}$$

$$= \left[\begin{array}{cc} 1 & 0 \\ -1/4 & -3/4 \\ \hline -3/4 & -1/4 \\ 0 & 1 \end{array}\right] \begin{bmatrix} \eta_1 \\ \eta_4 \end{bmatrix} =: \widehat{N}\widehat{w}$$

Check. We have

$$A'\widehat{N} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -3 & 1 & -2 \\ 1 & -2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/4 & -3/4 \\ -3/4 & -1/4 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = O$$

so, of course,

$$A'y = A'\widehat{N}\widehat{w} = O\widehat{w} = 0.$$

Here, neither the free variables nor the "basis matrices," N and $\widehat{N}$, are the same for the two pivot strategies. But *the matrices* N *and* $\widehat{N}$ *can be related.* Note that

$$Q'\widehat{N} = \begin{bmatrix} -1/4 & -3/4 \\ -3/4 & -1/4 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

has the same "trapezoidal" structure, $\begin{bmatrix} * \\ I_2 \end{bmatrix}$, as N. We can find matrices B and C so that

$$\mathrm{NB} = \widehat{\mathrm{N}} \quad \text{and} \quad \widehat{\mathrm{N}}\mathrm{C} = \mathrm{N}\,. \qquad (*)$$

This is done by comparing the last two rows of $\mathrm{NB} = \widehat{\mathrm{N}}$ and the first and last rows of $\widehat{\mathrm{N}}\mathrm{C} = \mathrm{N}$ or, equivalently, the last two rows of $\mathrm{Q}'\widehat{\mathrm{N}}\mathrm{C} = \mathrm{Q}'\mathrm{N}$. These rows involve the "identity parts" of N, $\widehat{\mathrm{N}}$ and $\mathrm{Q}'\widehat{\mathrm{N}}$, respectively. We find

$$\mathrm{B} = \begin{bmatrix} -3/4 & -1/4 \\ 0 & 1 \end{bmatrix}, \quad \mathrm{C} = \begin{bmatrix} -4/3 & -1/3 \\ 0 & 1 \end{bmatrix}.$$

The *full* relations $(*)$ are easily checked, as is the fact that

$$\mathrm{BC} = \mathrm{I}_2 = \mathrm{CB}\,.$$

Such relations, connecting "basis matrices" obtained via different pivot strategies will be seen to hold *in general.*

Problem 1. Do hand computations analogous with those on pages 12-21 for the matrix

$$\mathrm{A} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 5 \end{bmatrix}$$

and the two right sides

a. $$\mathrm{b} = \mathrm{Ax}_0 = \begin{bmatrix} 1 \\ 4 \\ 7 \\ 10 \end{bmatrix}, \quad \mathrm{x}_0 := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

b. $$\mathrm{b} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}.$$

You will get two *different* matrices M and $\widehat{\mathrm{M}}$, say. Relate these matrices as we related N and $\widehat{\mathrm{N}}$ above (M and $\widehat{\mathrm{M}}$ are one column matrices!).

Problem 2. Let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 3 \\ 1 & -1 & 4 \end{bmatrix}.$$

a. Compute $\rho(A)$.

b. Does the homogeneous system $Ax = 0$ have the unique solution $x = 0$?

c. If $Ax = b$ is solvable, is its solution unique?

d. Is $Ax = b$ always solvable (for *arbitrary* $b \in \mathbb{R}^4$)?

e. Does the homogeneous system $A'y = 0$ have the unique solution $y = 0$?

f. If $A'y = d$ is solvable, is its solution unique?

g. Is $A'y = d$ always solvable (for *arbitrary* $d \in \mathbb{R}^3$)?

h. Find the general solutions of $Ax = b$ for (i) $b = [1\ \ -1\ \ 1\ \ -1]'$, (ii) $b = [1\ \ 0\ \ 0\ \ 1]'$, (iii) $b = [4\ \ 3\ \ 2\ \ 1]'$.

i. Find the general solution of $A'y = d = [1\ \ 0\ \ 1]'$.

Nonsingular matrices.

We always have $A \in \mathbb{F}^{n \times m}$. We know that $\rho = \rho(A)$ depends only on A.

The matrix A represents a function (or transformation, or mapping) from $\mathbb{F}^m$ to $\mathbb{F}^n$. This function is given by

$$A : \mathbb{F}^m \to \mathbb{F}^n$$

$$x \to y := Ax\ ,$$

that is by *matrix-vector multiplication.*

Definitions.

$$\begin{aligned} \mathcal{R}(A) &:= \textit{range} \text{ of } A \\ &:= \{Ax : \quad x \in \mathbb{F}^m\} \\ &:= A\mathbb{F}^m. \end{aligned}$$

$$\mathcal{N}(a) := \textit{null space} \text{ of } A$$
$$:= \{x \in \mathbb{F}^m : \quad Ax = 0_n\}.$$

By definition, $\mathcal{R}(A)$ is the set of all vectors $b \in \mathbb{F}^n$ for which $Ax = b$ is solvable: $\mathcal{N}(A)$ is the set of all solutions of the homogeneous system $Ax = 0_n$.

Proposition 1.

1. $\mathcal{R}(A) = \mathbb{F}^n \Leftrightarrow \rho(A) = n$,
2. $\mathcal{N}(A) = \{0_m\} \Leftrightarrow \rho(A) = m$.

□

1. $\mathcal{R}(A) = \mathbb{F}^n$ *means that* $Ax = b$ is solvable for *every* $b \in \mathbb{F}^n$. If there is a solvability condition $L_2c = b_2$, that is if $\rho(A) < n$, we can always choose b_2 so it is not satisfied.
2. $\mathcal{N}(A) = \{0_m\}$ *means that* there are no free variables, that is, $\rho(A) = m$. ■

Definitions. The matrix $A \in \mathbb{F}^{n \times m}$ is *nonsingular* if, for *every* $b \in \mathbb{F}^n$, the linear system $Ax = b$ has a *unique* solution $x \in \mathbb{F}^n$. A is *singular* if it is not nonsingular.

Thus, A is singular if there is a vector $b \in \mathbb{F}^n$ so that $Ax = b$ is either not solvable, or it has more than one solution.

The solvability of $Ax = b$, for *every* $b \in \mathbb{F}^n$ means that $\rho(A) = n$. The *uniqueness* of solutions means that there are no free variables: $\rho(A) = m$. Thus we have

Proposition 2.

Only *square* matrices $(m = n)$ can be nonsingular. For $A \in \mathbb{F}^{n \times n}$ the following statements are equivalent:

1. A is nonsingular,
2. $\rho(A) = n$,
3. $\mathcal{N}(A) = \{0_n\}$,
4. $\mathcal{N}(A') = \{0_n\}$,
5. $\mathcal{R}(A) = \mathbb{F}^n$,
6. $\mathcal{R}(A') = \mathbb{F}^n$,

[7. $|\det A| = |\upsilon_{11} \ \upsilon_{22} \ \cdots \ \upsilon_{nn}| > 0$.]

Remarks.

Property 2 is the main *practical* test for nonsingularity (so far!). Property 3 is the main *theoretical* test for nonsingularity. To show that $A \in \mathbb{R}^{n\times n}$ is nonsingular, one typically shows that $Ax = 0 \Rightarrow x = 0$. We give examples of this below. On the other hand, $A \in \mathbb{R}^{n\times n}$ is singular if there is an $x \neq 0$ with $Ax = 0$. Properties 3 and 4 are equivalent since $\rho(A) = \rho(A')$. We downplay property 7, A nonsingular $\Leftrightarrow$ det $A \neq 0$, since det A is *extremely sensitive* to small changes in A, even for moderate n. We shall later discuss more "robust" measures (cond A) of the "near singularity" of A. It *is* useful to note that det A is $(-1)^{\upsilon}$ times the *product of the pivots*, where υ is the total number of interchanges (of columns and/or of rows) made in the factorization process.

Examples.

1. Permutation matrices P are nonsingular. Since $P'P = I$ then $Px = b \Rightarrow x = P'b$ (uniqueness of solutions). Then, since $PP' = I$, $x = P'b \Rightarrow Px = b$ (existence of solutions).

2. Lower or upper triangular matrices are nonsingular if and only if all their diagonal elements are not zero. Then, e.g., $Lx = b$ can be f-solved uniquely for x. If some diagonal element is zero, we run into an equation $0\,\xi_j = \gamma_j$. If $\gamma_j \neq 0$, this has *no solution*. If $\gamma_j = 0$, it has *infinitely many solutions*.

3. $W_5 = \text{mxwilkinson}(5)$

$$= \begin{bmatrix} 1 & & & & 1 \\ -1 & 1 & & & 1 \\ -1 & -1 & 1 & & 1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ -1 & -1 & 1 & & \\ -1 & -1 & -1 & & \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & 2 \\ & & 1 & & 4 \\ & & & 1 & 8 \\ & & & & 16 \end{bmatrix}$$

   is nonsingular [with det $W_5 = 2^{5-1}$; in general det $W_n = 2^{n-1}$].

4. $T_5 = \text{mxtsd}(5)$

$$= \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ -1/2 & 1 & & & \\ & -2/3 & 1 & & \\ & & -3/4 & 1 & \\ & & & -4/5 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & & & \\ & 3/2 & -1 & & \\ & & 4/3 & -1 & \\ & & & 5/4 & -1 \\ & & & & 6/5 \end{bmatrix}$$

$$
= \begin{bmatrix} 1 & & & & \\ -1/2 & 1 & & & \\ & -2/3 & 1 & & \\ & & -3/4 & 1 & \\ & & & -4/5 & 1 \end{bmatrix}
\begin{bmatrix} 2 & & & & \\ & 3/2 & & & \\ & & 4/3 & & \\ & & & 5/4 & \\ & & & & 6/5 \end{bmatrix}
\begin{bmatrix} 1 & -1/2 & & & \\ & 1 & -2/3 & & \\ & & 1 & -3/4 & \\ & & & 1 & -4/5 \\ & & & & 1 \end{bmatrix}
$$

$$= \mathrm{LDL}'$$

is nonsingular [with det $T_5 = 6$; in general det $T_n = n + 1$].

5. $P_5 = \mathrm{mxpsd}(5)$

$$
= \begin{bmatrix} 2 & -1 & & & -1 \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ -1 & & & -1 & 2 \end{bmatrix}
= \begin{bmatrix} 1 & & & & \\ -1/2 & 1 & & & \\ & -2/3 & 1 & & \\ & & -3/4 & 1 & \\ -1/2 & -1/3 & -1/4 & -1 & 1 \end{bmatrix}
\begin{bmatrix} 2 & -1 & & & -1 \\ & 3/2 & -1 & & -1/2 \\ & & 4/3 & -1 & -1/3 \\ & & & 5/4 & -5/4 \\ & & & & 0 \end{bmatrix}
$$

$$
= \left[\begin{array}{cccc|c} 1 & & & & \\ -1/2 & 1 & & & \\ & -2/3 & 1 & & \\ & & -3/4 & 1 & \\ -1/2 & -1/3 & -1/4 & - & 1 \end{array}\right]
\left[\begin{array}{cccc|c} 2 & & & & \\ & 3/2 & & & \\ & & 4/3 & & \\ & & & 5/4 & \\ \hline & & & & 0 \end{array}\right]
\left[\begin{array}{ccccc} 1 & -1/2 & & & -1/2 \\ & 1 & -2/3 & & -1/3 \\ & & 1 & -3/4 & -1/4 \\ & & & 1 & -1 \\ \hline & & & & 1 \end{array}\right]
$$

$$
= \begin{bmatrix} 1 & & & \\ -1/2 & 1 & & \\ & -2/3 & 1 & \\ & & -3/4 & 1 \\ -1/2 & -1/3 & -1/4 & -1 \end{bmatrix}
\begin{bmatrix} 2 & & & \\ & 3/2 & & \\ & & 4/3 & \\ & & & 5/4 \end{bmatrix}
\begin{bmatrix} 1 & -1/2 & & & -1/2 \\ & 1 & -2/3 & & -1/3 \\ & & 1 & -3/4 & -1/4 \\ & & & 1 & -1 \end{bmatrix}
$$

$$= \mathrm{LDL}'$$

is singular [with det $P_n \equiv 0$, $n \geq 3$].

6. $M_5 = \text{mxmin}(5)$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 1 \\ & & 1 & 1 & 1 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}$$

$= LL'$

is nonsingular [with det $M_n \equiv 1$].

7. $M_5 = \text{mxmax}(5)$

$$= \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 3 & 4 & 5 \\ 3 & 3 & 3 & 4 & 5 \\ 4 & 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 & 5 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ 2 & 1 & & & \\ 3 & 3/2 & 1 & & \\ 4 & 4/2 & 4/3 & 1 & \\ 5 & 5/2 & 5/3 & 5/4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ & -2 & -3 & -4 & -5 \\ & & -3/2 & -4/2 & -5/2 \\ & & & -4/3 & -5/3 \\ & & & & -5/4 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & & & & \\ 2 & 1 & & & \\ 3 & 3/2 & 1 & & \\ 4 & 4/2 & 4/3 & 1 & \\ 5 & 5/2 & 5/3 & 5/4 & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & -2 & & & \\ & & -3/2 & & \\ & & & -4/3 & \\ & & & & -5/4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ & 1 & 3/2 & 4/2 & 5/2 \\ & & 1 & 4/3 & 5/3 \\ & & & 1 & 5/4 \\ & & & & 1 \end{bmatrix}$$

$= LDL'$

is nonsingular [with det $M_n = (-1)^{n-1}n$].

8. $M_5 = \text{mxmix}(5)$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 2 & 3 & 2 & 1 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \left[\begin{array}{ccc|cc} 1 & & & & 0 \\ 1 & 1 & & & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & & & 0 \\ 1 & & & & 0 \end{array}\right] \left[\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & \\ & & 1 & & \\ \hline & & 0 & & \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$= \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 1 & 1 & 1 \\ 1 & 1 & \\ 1 & & \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & \\ & & 1 & & \end{bmatrix} = LL'$$

is singular [with $\det M_n \equiv 0$, $n > 1$].

9. The special *arrow matrix*

$$A_n = \text{mxarrow}(n)$$
$$:= \begin{bmatrix} I & e \\ e' & 1 \end{bmatrix}, \qquad e := \text{ones}(n-1,\ 1)$$

has the LU factorization

$$A_n = \begin{bmatrix} I & \\ e' & 1 \end{bmatrix} \begin{bmatrix} I & e \\ & -m \end{bmatrix}, \qquad m := n-2$$
$$= \begin{bmatrix} I & \\ e' & 1 \end{bmatrix} \begin{bmatrix} I & \\ & -m \end{bmatrix} \begin{bmatrix} I & e \\ & 1 \end{bmatrix}$$
$$= LDL'.$$

Thus A_n is nonsingular if and only if $n \neq 2$ [$\det A_n = 2 - n$].

For general matrices we *must* use a complete pivoting strategy (at least for "nonzeroness," if not for "size"). This is moderately expensive. Partial pivoting is substantially cheaper.

Proposition 3.

For nonsingular matrices A, partial pivoting suffices (to give $\rho(A) = n$ pivots).

□

1. There's a pivot in the first column. Otherwise, $Ae_1 = 0$, $e_1 := [1\ \ 0\ \ 0\ \ \ldots\ \ 0]'$, and A would be singular (proposition 2.2).

2. If Q is a permutation matrix then $Q'A$ is nonsingular. For if $Q'Ax = 0$ then $Ax = QQ'Ax = 0$ so $x = 0$ since A is nonsingular. Again we used proposition 2.2, twice.

3. We now factor

$$Q'A = \begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix} = \begin{bmatrix} 1 & \\ \ell & I \end{bmatrix} \begin{bmatrix} \alpha & b' \\ & G \end{bmatrix} \qquad (\alpha \neq 0).$$

The *key step* is to show that $\begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix}$ nonsingular (and $\alpha \neq 0$) $\Rightarrow$ G nonsingular.

Suppose $Gx = 0$ to show that $x = 0$.

Define ξ by

$$\alpha\xi + b'x = 0 \,,$$

that is

$$\xi := -\frac{b'x}{\alpha} \,.$$

Then

$$\begin{bmatrix} \alpha & b' \\ & G \end{bmatrix} \begin{bmatrix} \xi \\ x \end{bmatrix} = 0$$

and multiplication by $\begin{bmatrix} 1 & \\ \ell & I \end{bmatrix}$ gives

$$\underbrace{\begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix}}_{Q'A} \begin{bmatrix} \xi \\ x \end{bmatrix} = 0 \,.$$

Since $Q'A$ is nonsingular, then $\begin{bmatrix} \xi \\ x \end{bmatrix} = 0$; in particular, $x = 0$. Thus G is nonsingular.

4. Reduction!

■

Definition. GL(n) is the set of nonsingular $n \times n$ matrices. (This stands for *general linear "group."* What follow are the "group properties" of the nonsingular $n \times n$ matrices.)

Proposition 4.

Let A, $B \in \mathbb{F}^{n \times n}$. Then:

1. A, $B \in GL(n) \;\Rightarrow\; AB \in GL(n)$.
2. $AB \in GL(n) \;\Rightarrow\;$ A, $B \in GL(n)$.

3. $A \in GL(n) \Rightarrow$ there's a unique $B \in GL(n)$ with $AB = I_n$.

4. $AB = I_n \Rightarrow BA = I_n$ (*one side is enough*).

Definition. If

$$AB = I_n = BA$$

we say that A and B are *inverses* of each other and write

$$A^{-1} := B, \quad B^{-1} := A .$$

5. $A, B \in GL(n) \Rightarrow (AB)^{-1} = B^{-1}A^{-1}$.

6. $A \in GL(n) \Leftrightarrow A' \in GL(n)$, and $(A')^{-1} = (A^{-1})' =: A^{-\prime}$.

□

1. This might be called the *factorization principle.* We must show that, for every given $b \in \mathbb{F}^n$, the linear system $ABx = b$ has a unique solution $x \in \mathbb{F}^n$. Well, let $c := Bx$. Then $ABx = b \Leftrightarrow Ac = b$ and $Bx = c$. If $b \in \mathbb{F}^n$ is given then c is uniquely determined, since $A \in GL(n)$. Then x is uniquely determined since $B \in GL(n)$.

2. This is the *converse* of 1. It is a little deeper. Let $AB \in GL(n)$. Let $Bx = 0$ to show that $x = 0$. Then also $ABx = 0$, so $x = 0$ since $AB \in GL(n)$. Thus $B \in GL(n)$. That was easy, but what about A? Well, we know know that $A \in GL(n) \Leftrightarrow A' \in GL(n)$ (proposition 2.4). So $AB \in GL(n) \Leftrightarrow (AB)' = B'A' \in GL(n)$ (reverse order rule for *transposition*). Thus, as above, $A' \in GL(n)$, that is, $A \in GL(n)$.

3. $AB = I_n \Leftrightarrow Ab_i = e_i$, $1 \le i \le n$, where the b_i and e_i are the columns of B and I, respectively. Since $A \in GL(n)$ all these equations have unique solutions. *There's more!* Since, clearly, $AB = I_n \in GL(n)$ then 2 shows that $B \in GL(n)$, *as required.*

4. Suppose $AB = I_n$. Again $B \in GL(n)$ by 2. By 3 there's a $C \in GL(n)$ with $BC = I_n$. We then have

$$BA = BAI_n = BABC = BI_nC = BC = I_n ,$$

as required.

5. Let $A, B \in GL(n)$. Then $AB \in GL(n)$ by 1. By 3 and 4 $(AB)^{-1}$ exists and is unique. But clearly $(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = I_n$, so $(AB)^{-1} = B^{-1}A^{-1}$ by the uniqueness.

6. We already know the stated equivalence (propositions 2.3, 2.4). We have $A^{-1}A = I_n$. By the reverse order rule for *transposition*, $A'(A^{-1})' = I_n$. Thus $(A')^{-1} = (A^{-1})'$ by the uniqueness.

■

Example.

Let $M = \text{mxmax}(5)$

$$= \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 3 & 4 & 5 \\ 3 & 3 & 3 & 4 & 5 \\ 4 & 4 & 4 & 4 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{bmatrix}.$$

$M = M'$ is real symmetric.

In general

$$J := J_n := \begin{bmatrix} e_n \, e_{n-1} \cdots \; e_1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (n = 5)$$

is the ***counter identity*** and, for $A \in \mathbb{C}^{n \times n}$,

$$\text{flip } A := JA^T J$$

is the "***counterpose,***" or simply the "flip," of A. flip A is the reflection of A in its ***counter diagonal*** ( ). A is ***counter symmetric*** if $A = \text{flip } A$.

Since $M = M'$ is real,

flip $M = JMJ$

$$= \begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ 5 & 4 & 4 & 4 & 4 \\ 5 & 4 & 3 & 3 & 3 \\ 5 & 4 & 3 & 2 & 2 \\ 5 & 4 & 3 & 2 & 1 \end{bmatrix}$$

$$
= \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ & -1 & -1 & -1 & -1 \\ & & -1 & -1 & -1 \\ & & & -1 & -1 \\ & & & & -1 \end{bmatrix}
$$

$$
= \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & & & & \\ & -1 & & & \\ & & -1 & & \\ & & & -1 & \\ & & & & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 1 \\ & & 1 & 1 & 1 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}
$$

$=: \mathrm{LDL}'$.

Now $\mathrm{J} = \mathrm{J}'$ is a permutation matrix so $\mathrm{J}^2 = \mathrm{I}$, that is $\mathrm{J}^{-1} = \mathrm{J}$. Thus

$$
\mathrm{M} = \mathrm{JLDL'J}\,,
$$

being a product of (five) nonsingular matrices, is also nonsingular (proposition 4.1). Now also

$$
\mathrm{M} = (\mathrm{JL})\,\mathrm{D}\,(\mathrm{JL})'.
$$

The matrix JL is L "turned upside down," and *it is symmetric.*

Thus

$\mathrm{M} = (\mathrm{JL})\,\mathrm{D}\,(\mathrm{JL})$

$$
= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & \\ 1 & 1 & 1 & & \\ 1 & 1 & & & \\ 1 & & & & \end{bmatrix} \begin{bmatrix} 5 & & & & \\ & -1 & & & \\ & & -1 & & \\ & & & -1 & \\ & & & & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & \\ 1 & 1 & 1 & & \\ 1 & 1 & & & \\ 1 & & & & \end{bmatrix}
$$

Let us compute

$$
\mathrm{M}^{-1} = \mathrm{JL}^{-\prime}\mathrm{D}^{-1}\mathrm{L}^{-1}\mathrm{J}\ .
$$

For this we need L^{-1}.

In general,

$$E := E_n := \left[\begin{array}{ccccc} e_2 & e_3 & \dots & e_n & 0 \end{array}\right]$$

$$= \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \quad (n = 5)$$

is the *downshift matrix* (of order n). We have

$$Ex = E\left[\begin{array}{cccc} \xi_1 & \xi_2 & \dots & \xi_n \end{array}\right]'$$

$$= \left[\begin{array}{ccccc} 0 & \xi_1 & \xi_2 & \dots & \xi_{n-1} \end{array}\right]'.$$

We have $(n = 5)$

$$E^2 = \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right], \quad E^3 = \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right],$$

$$E^4 = \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right], \quad E^5 = O.$$

Thus, for *general* n, the geometric series

$$I + E + E^2 + \dots = I + E + E^2 + \dots + E^{n-1}$$

"converges" and equals L. Let

$$B := I - E.$$

Then

$$\begin{aligned} BL &= (I-E)(I+E+E^2+\dots) \\ &= I+E+E^2+E^3+\dots \\ &\quad -E-E^2-E^3-\dots \\ &= I = LB \end{aligned}$$

and so

$$L^{-1} = B = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{bmatrix} \quad (n=5).$$

We now have

$JM^{-1}J = B'D^{-1}B$

$$= \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & -1 & \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix} \begin{bmatrix} 1/5 & & & & \\ & -1 & & & \\ & & -1 & & \\ & & & -1 & \\ & & & & -1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & -1 & \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix} \begin{bmatrix} 1/5 & & & & \\ 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & -1 & \\ & & & 1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -4/5 & 1 & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{bmatrix} = - \begin{bmatrix} 4/5 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}$$

Thus M^{-1} is the "flip" of this real symmetric matrix:

$$M^{-1} = -\begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 4/5 \end{bmatrix} \quad (n=5)$$

$$= -\begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & 2 & -1 \\ & & & -1 & 1-\frac{1}{n} \end{bmatrix}, \quad \text{in general,}$$

$$= -\begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} + \frac{1}{n} e_n e_n' .$$

Some uses of inverses.

Block Gauss factorization.

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{11} \text{ nonsingular,}$$

$$= \begin{bmatrix} I & \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ & G \end{bmatrix}$$

with

$$G = A_{22} - A_{21} A_{11}^{-1} A_{12}$$

the *block Gauss transform* or, more precisely, the *Schur complement* of the *pivot block* A_{11} in A. [If A is also square then

$$\det A = \det A_{11} \det\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)$$

(*Sylvester's determinant identity*).

In particular,

$$\det\begin{bmatrix} \alpha & b' \\ a & B \end{bmatrix} = \alpha \det\left(B - \frac{ab'}{\alpha}\right)$$

gives a *better way* to compute determinants than by the Laplace expansion, the method that is usually taught. In fact this is just Gauss factorization!

Examples.

1. For $A = \begin{bmatrix} I & e \\ e' & 1 \end{bmatrix}$ the $n \times n$ *arrow matrix* we have

$$\det A = \det I\,(1 - e'e)$$
$$= 1 - (n-1) = 2 - n.$$

2. *In general,* $\lambda \in \mathbf{C}$ is an *eigenvalue* of $A \in \mathbf{C}^{n\times n}$ if $\lambda I - A$ is singular. Thus the eigenvalues of A are the zeros of the *characteristic polynomial* $p(\lambda) := p_A(\lambda) := \det(\lambda I - A)$. For our *arrow matrix* A, Sylvester's determinant identity gives

$$p(\lambda) = \det\begin{bmatrix} \lambda I - I & -e \\ -e' & \lambda - 1 \end{bmatrix}$$

$$= \det\,[(\lambda-1)\,I]\left(\lambda - 1 - \frac{e'e}{\lambda - 1}\right)$$

$$= (\lambda-1)^{n-1}\left(\lambda - 1 - \frac{n-1}{\lambda-1}\right)$$

$$= (\lambda-1)^{n-2}\left[(\lambda-1)^2 - (n-1)\right]$$

$$= (\lambda-1)^{n-2}\left(\lambda^2 - 2\lambda - m\right), \qquad m := n-2$$

Thus A has $n-2$ eigenvalues $\lambda_k \equiv 1$ $(1 < k < n)$. The other two eigenvalues solve

$$\lambda^2 - 2\lambda - m = 0$$

and are thus

$$\lambda_1 := \sqrt{n-1} + 1$$
$$\sim \sqrt{n}, \qquad n \to +\infty,$$

and

$$\lambda_n := -\frac{n-2}{\sqrt{n-1}+1}$$

$$\sim -\sqrt{n}, \quad n \to +\infty.]$$

Sherman-Morrison-Woodbury formula: if S and T are nonsingular and $\boxed{S+T=U'V}$ then

$$\boxed{(I-VS^{-1}U')(I-VT^{-1}U')=I}\,.$$

□

$$\begin{aligned}(I-VS^{-1}U')(I-VT^{-1}U') \\ &= I-VS^{-1}U'-VT^{-1}U'+VS^{-1}U'VT^{-1}U' \\ &= I-VS^{-1}(S+T-U'V)\,T^{-1}U'=I\,.\end{aligned}$$

■

Sherman-Morrison formula ($S=\sigma=-1$). If $1+u'v\neq 0$, then

$$\boxed{(I+vu')^{-1}=I-\frac{vu'}{1+u'v}}\,.$$

If $u'v=-1$ then $I+vu'$ is singular, since $v\neq 0$ and

$$\begin{aligned}(I+vu')v &= v+vu'v \\ &= v+v(-1) \\ &= 0.\end{aligned}$$

[By Sylvester's identity, and some interchanges which leave the determinant unchanged,

$$\begin{aligned}1-u'v = det\begin{bmatrix} I & v \\ u' & 1\end{bmatrix} &= det\begin{bmatrix} 1 & u' \\ V & I\end{bmatrix} \\ &= det\,(I-vu')\,.\end{aligned}$$

That is ($v\leftrightarrow -v$).

$$\boxed{det\,(I+vu')=1+u'v}\,.\,]$$

Some explicit inverses.

1. Special arrow matrix

$$A=\begin{bmatrix} I & e \\ e' & 1\end{bmatrix}, \qquad e=\text{ones}\,(n-1,\,1)$$

$$= \begin{bmatrix} I & \\ e' & 1 \end{bmatrix} \begin{bmatrix} I & \\ & -m \end{bmatrix} \begin{bmatrix} I & e \\ & 1 \end{bmatrix}, \qquad m := n-2 \neq 0 .$$

General fact:

$$\begin{bmatrix} I & \\ L & I \end{bmatrix}^{-1} = \begin{bmatrix} I & \\ -L & I \end{bmatrix}, \quad \begin{bmatrix} I & U \\ & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -U \\ & I \end{bmatrix},$$

as can easily be checked by block multiplication.

By this and the reverse order rule for matrix inversion,

$$A^{-1} = \begin{bmatrix} I & -e \\ & 1 \end{bmatrix} \begin{bmatrix} I & \\ & -1/m \end{bmatrix} \begin{bmatrix} I & \\ -e' & 1 \end{bmatrix}$$

$$= \frac{1}{m} \begin{bmatrix} I & -e \\ & 1 \end{bmatrix} \begin{bmatrix} mI & \\ & -1 \end{bmatrix} \begin{bmatrix} I & \\ -e' & 1 \end{bmatrix}$$

$$= \frac{1}{m} \begin{bmatrix} I & -e \\ & 1 \end{bmatrix} \begin{bmatrix} mI & \\ e' & -1 \end{bmatrix}$$

$$= \frac{1}{m} \begin{bmatrix} mI - ee' & e \\ e' & -1 \end{bmatrix}$$

$$= \frac{1}{3} \left[\begin{array}{cccc|c} 2 & -1 & -1 & -1 & 1 \\ -1 & 2 & -1 & -1 & 1 \\ -1 & -1 & 2 & -1 & 1 \\ -1 & -1 & -1 & 2 & 1 \\ \hline 1 & 1 & 1 & 1 & -1 \end{array} \right] \qquad (n=5) .$$

Our matlab functions for A and the matrix mA^{-1}, whose elements are *integers*, are mxarrow and mxworra, respectively.

2. The "min matrix"

$$M = \text{mxmin}(n)$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \quad (n = 5)$$

$$= \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 1 \\ & & 1 & 1 & 1 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix} =: LL'.$$

We have already seen (page 33) that $L^{-1} = B = I - E$. Thus

$$M^{-1} = B'B = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & -1 & \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} \quad (n = 5)$$

$$= T - e_n e_n', \text{ in general,}$$

where

$$T := \text{mxtsd}(n)$$

is the ***negative second difference matrix.***

3. The negative second difference matrix

$T := \mathrm{mxtsd}(n)$

$$= \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \quad (n=5)\,.$$

Continuing #2 we have

$$\begin{aligned} T &= B'B + e_n e_n' \\ &= B'\left(I + B^{-\prime} e_n e_n' B^{-1}\right)B \\ &= B'(I + L' e_n e_n' L)\,B \\ &= B'(I + ee')\,B \end{aligned}$$

with

$$e := \mathrm{ones}(n,\ 1)\,.$$

By the Sherman-Morrison formula,

$$\begin{aligned} T^{-1} &= B^{-1}\,(I + ee')^{-1} B^{-1} \\ &= L\left(I - \frac{ee'}{1+e'e}\right)L' \\ &= LL' - \frac{Le(Le)'}{n+1} \\ &= M - \frac{uu'}{n+1} \end{aligned}$$

with

$$u := Le = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \qquad n=5,$$

$$= \begin{bmatrix} 1 & 2 & 3 & \dots & n \end{bmatrix}', \qquad \text{in general.}$$

Thus, for n = 5,

$$T^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix}.$$

That is

$$6T^{-1} = \begin{bmatrix} 6 & 6 & 6 & 6 & 6 \\ 6 & 12 & 12 & 12 & 12 \\ 6 & 12 & 18 & 18 & 18 \\ 6 & 12 & 18 & 24 & 24 \\ 6 & 12 & 18 & 24 & 30 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \\ 3 & 6 & 9 & 12 & 15 \\ 4 & 8 & 12 & 16 & 20 \\ 5 & 10 & 15 & 20 & 25 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & 8 & 6 & 4 & 2 \\ 3 & 6 & 9 & 6 & 3 \\ 2 & 4 & 6 & 8 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

The *general pattern* is easy to see. We have

$$G := \begin{bmatrix} \gamma_{ji} \end{bmatrix} := (n+1)T$$

with

$$\begin{aligned} \gamma_{ji} &= \beta_j \alpha_i , \qquad i \le j , \\ &= \alpha_j \beta_i , \qquad i \ge j , \end{aligned} \tag{$*$}$$

and

$$a := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} := u = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix},$$

$$b := \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} := \begin{bmatrix} n \\ n-1 \\ \vdots \\ 1 \end{bmatrix} = Ja\,.$$

A matrix $G = \left[\, \gamma_{ji} \,\right]$ whose elements are of the form (*), with *arbitrary* vectors a, $b \in \mathbb{R}^n$, is said to be a *Green's matrix*. Thus, Green's matrices are symmetric and they depend on at most 2n real numbers. Our *special* Green's matrix is also counter symmetric.

4. *Wilkinson's "pivoting matrix,"*

W = mxwilkinson(n)

$$= \begin{bmatrix} 1 & & & & 1 \\ -1 & 1 & & & 1 \\ -1 & -1 & 1 & & 1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix} \quad (n=5)$$

$$= \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ -1 & -1 & 1 & & \\ -1 & -1 & -1 & 1 & \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 16 \end{bmatrix} \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & 2 \\ & & 1 & & 4 \\ & & & 1 & 8 \\ & & & & 1 \end{bmatrix}$$

$=:$ LDU.

We *know* that U is easy to invert. We must invert

$$L = I - E - E^2 - E^3 - \ldots .$$

We know that $E^n = 0$. By the geometric series, for $n = 5$,

$$\begin{aligned} L^{-1} = I &+ \left(E + E^2 + E^3 + E^4\right) \\ &+ \left(E + E^2 + E^3\right)^2 \\ &+ \left(E + E^2\right)^3 \\ &+ E^4 \end{aligned}$$

$$
\begin{aligned}
&= I + E + E^2 + \ E^3 + \ E^4 \\
&\qquad\quad + E^2 + 2E^3 + 3E^4 \\
&\qquad\qquad\quad + \ E^3 + 3E^4 \\
&\qquad\qquad\qquad\quad + \ E^4 \\
&= I + E + 2E^2 + 4E^3 + 8E^4
\end{aligned}
$$

$$
= \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 2 & 1 & 1 & & \\ 4 & 2 & 1 & 1 & \\ 8 & 4 & 2 & 1 & 1 \end{bmatrix}.
$$

The *general* pattern is clear. L^{-1} is unit lower triangular with the constant value 2^k along its kth *subdiagonal*. For $n = 5$,

$W^{-1} =$

$$
\begin{bmatrix} 1 & & & & -1 \\ & 1 & & & -2 \\ & & 1 & & -4 \\ & & & 1 & -8 \\ & & & & 1 \end{bmatrix}
\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1/16 \end{bmatrix}
\begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 2 & 1 & 1 & & \\ 4 & 2 & 1 & 1 & \\ 8 & 4 & 2 & 1 & 1 \end{bmatrix},
$$

that is

$2^{n-1} W^{-1} =$

$$
\begin{bmatrix} 16 & & & & -1 \\ & 16 & & & -2 \\ & & 16 & & -4 \\ & & & 16 & -8 \\ & & & & 1 \end{bmatrix}
\begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 2 & 1 & 1 & & \\ 4 & 2 & 1 & 1 & \\ 8 & 4 & 2 & 1 & 1 \end{bmatrix}
$$

$$
= \begin{bmatrix} 8 & -4 & -2 & -1 & -1 \\ 0 & 8 & -4 & -2 & -2 \\ 0 & 0 & 8 & -4 & -4 \\ 0 & 0 & 0 & 8 & -8 \\ 8 & 4 & 2 & 1 & 1 \end{bmatrix} .
$$

Again, the pattern for *general* n should be clear. Our code for $2^{n-1}W^{-1}$ is mxnosnikliw.

Problems

These are mainly experimental problems, using matlab. But, what can you prove?

Problem 3.

For general n, what is the largest value γ_{ji} in our special Green's matrix $G = G_n$? For precisely which index pairs (i, j) is the maximum attained? (No proof required. Experiment, using the code mxgreen. Try "meshing" G_n for large n, n = 200 say (help mesh).

Problem 4.

For general n, what are L_n and D_n in the factorization $G = L_n D_n L_n'$: $L_n =$, $D_n =$. (Again, no formal proof is required, just the correct answer! See matlab's rat function (help rat). I get

$$
\text{diag } D_n = 9\left[\frac{8}{9} \ \ \frac{7}{8} \ \ \frac{6}{7} \ \ \frac{5}{6} \ \ \frac{4}{5} \ \ \frac{3}{4} \ \ \frac{2}{3} \ \ \frac{1}{2}\right]
$$

for n = 8. The pattern is clear.)

Problem 5.

What are the lower triangular matrices L_n in the LU factorizations of $A_n :=$ mxnosnikliw(n)? What are their inverses L_n^{-1}?

Problem 6.

Gauss factorization with no pivoting breaks down on the nonsingular matrices $B_n :=$ mxworra(n), for $n \geq 3$. At which stage does there occur a zero in the pivot position, as a function of n (for n = 3, 4, ..., 10, say)?

Equivalent matrices.

Definition. Two matrices $A, B \in \mathbb{F}^{n \times m}$ are *equivalent* ($A \sim B$) if there are nonsingular matrices F and G with

$$
AF = GB .
$$

Then, with the *changes of variables*

$$x :=: Fz \quad \text{and} \quad b :=: Gc$$

in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, the linear systems

$$Ax = b \quad \text{and} \quad Bz = c$$

are equivalent.

□

$$\begin{aligned} Ax = b &\Leftrightarrow AFz = Gc \\ &\Leftrightarrow GBz = Gc \\ &\Leftrightarrow Bz = c, \end{aligned}$$

since G is nonsingular.

■

Basic fact GF4. Matrix equivalence is an *equivalence* relation. That is it is:

1. *reflexive:* $A \sim A$
2. *symmetric:* $A \sim B \Leftrightarrow B \sim A$
3. *transitive:* $A \sim B$ and $B \sim C \Rightarrow A \sim C$.

□

1. Take $F = I_m$, $G = I_n$.
2. If $AF = GB$ with $F \in GL(m)$ and $G \in GL(m)$, then $BF^{-1} = G^{-1}A$ with $F^{-1} \in GL(m)$ and $G^{-1} \in GL(n)$, and vice versa.
3. If $AF = GB$ and $BH = KC$ with $F, H \in GL(m)$ and $G, K \in GL(n)$, then $AFH = GKC$ with $FH \in GL(m)$ and $GK \in GL(n)$.

■

We wish to chose F and G so that B is as simple as possible.

Proposition 5. Let $O \neq A \in \mathbb{F}^{n \times m}$.

1. There are nonsingular matrices F and G with

$$AF = GE_\rho: \quad E_\rho := \begin{bmatrix} I_\rho & O \\ O & O \end{bmatrix}.$$

2. In this decomposition ρ is unique. In fact $\rho = \rho(A)$, the *common* number of pivots in *any* Gauss factorization of A.

□

1. Let

$$Q'AP = LU = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix}$$

be an LU factorization. Then also

$$Q'AP = \begin{bmatrix} L_1 & \\ L_1 & I \end{bmatrix} \underbrace{\begin{bmatrix} I_\rho & O \\ O & O \end{bmatrix}}_{E_\rho} \begin{bmatrix} U_1 & U_2 \\ & I \end{bmatrix}.$$

Let $\begin{bmatrix} V_1 & V_2 \\ & I \end{bmatrix}$ solve

$$\begin{bmatrix} U_1 & U_2 \\ & I \end{bmatrix} \begin{bmatrix} V_1 & V_2 \\ & I \end{bmatrix} = I ,$$

that is

$$U_1 V_1 = I_\rho, \qquad U_1 V_2 + U_2 = O .$$

With

$$U = \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \dots & u_m \end{bmatrix}$$

$$\begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \dots & v_m \end{bmatrix}$$

these equations are equivalent with the upper triangular systems

$$U_1 v_i = e_i , \qquad 1 \le i \le \rho ,$$

$$U_1 v_i = -u_i , \qquad \rho < i \le m , ,$$

which, because $U_1 = \nabla$ is nonsingular, can be backsolved uniquely for $v_1, v_2, \dots, v_m$. In fact, $V_1 = \nabla$ is also nonsingular, its diagonal elements being the reciprocals of those of U_1.

We now have

$$\underbrace{A\,P \begin{bmatrix} V_1 & V_2 \\ & I \end{bmatrix}}_{=:\,F} = \underbrace{Q \begin{bmatrix} L_1 & \\ L_2 & I \end{bmatrix}}_{=:\,G} E_\rho$$

with F and G nonsingular. [In fact we have, *formally*,

$$V_1 = U_1^{-1}, \qquad V_2 = -U_1^{-1}U_2 = -V_1U_2.]$$

2. Suppose $A \sim E_\rho$ and $A \sim E_\sigma$. Then $E_\rho \sim E_\sigma$. We wish to show that $\rho = \sigma$. Suppose $\rho \neq \sigma$ to obtain a contradiction. By interchanging the roles of ρ and σ, if necessary, we may suppose that $\rho > \sigma$. Let $E_\rho F = GE_\sigma$ with F and G nonsingular. Partition

$$F = \begin{array}{cc} \left[\begin{array}{cc} F_{11} & F_{12} \\ F_{21} & F_{22} \end{array}\right] & \begin{array}{c} \rho \\ m-\rho \end{array} \\ \begin{array}{cc} \sigma & m-\sigma \end{array} & \end{array}$$

and

$$G = \begin{array}{cc} \left[\begin{array}{cc} G_1 & G_2 \end{array}\right] & n \\ \begin{array}{cc} \sigma & n-\sigma \end{array} & \end{array}.$$

Then

$$\begin{bmatrix} F_{11} & F_{12} \\ O & O \end{bmatrix} = \begin{bmatrix} I_\rho & O \\ O & O \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$$

$$= E_\rho F = GE_\sigma$$

$$= \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} I_\sigma & O \\ O & \end{bmatrix}$$

$$= \quad G_1 \quad [O]$$

so

$$F_{12} = O, \qquad F = \begin{bmatrix} F_{11} & O \\ F_{21} & F_{22} \end{bmatrix}$$

with $F_{22} \in \mathbb{F}^{(m-\rho)\times(m-\sigma)}$ "short and fat." By key fact GF2, there is an $x_2 \neq 0$ with $F_{22}x_2 = 0$.

But then $Fx = 0$ with $x := \begin{bmatrix} 0_\sigma \\ x_2 \end{bmatrix} \neq 0$, contradicting the nonsingularity of F. ■

Definition. The matrix E_ρ of proposition 5 is called *the canonical form* of A under matrix equivalence.

The word "canonical" connotes *uniqueness*. The *only invariant* of a matrix A under matrix equivalence is $\rho = \rho(A)$.

Thus, with nonsingular matrices F and G of proposition 5, and the changes of variables

$$x :=: Fz, \qquad b :=: Gc,$$

the linear system

$$Ax = b$$

becomes extremely simple:

$$E_\rho z = c.$$

Examples.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix}.$$

1. *No pivoting.*

$$A = LU = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \left[\begin{array}{cc} 1 & \\ 1 & 1 \\ \hline 1 & -1/3 \\ 1 & 2/3 \end{array}\right] \left[\begin{array}{cc|c} 1 & 0 & 1 \\ & -3 & -3 \end{array}\right].$$

$$G = \begin{bmatrix} L_1 & \\ L_2 & I_2 \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & & \\ 1 & 1 & & \\ \hline 1 & -1/3 & 1 & \\ 1 & 2/3 & 0 & 1 \end{array}\right].$$

Solve $U_1 V_1 = I_2$ for V_1:

$$\begin{bmatrix} 1 & \\ & -3 \end{bmatrix} V_1 = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}, \qquad V_1 = \begin{bmatrix} 1 & \\ & -1/3 \end{bmatrix}.$$

Solve $U_1 V_2 + U_2 = O$ for V_2:

$$\begin{bmatrix} 1 & \\ & -3 \end{bmatrix} V_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \qquad V_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

Then

$$F = \begin{bmatrix} V_1 & V_2 \\ & I_1 \end{bmatrix} = \left[\begin{array}{cc|c} 1 & & -1 \\ & -1/3 & -1 \\ \hline & & 1 \end{array}\right].$$

2. Complete pivoting.

$$Q'AP = LU = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \left[\begin{array}{cc} 1 & \\ -1/3 & 1 \\ \hline 0 & 3/4 \\ 2/3 & 1/4 \end{array}\right] \left[\begin{array}{cc|c} -3 & 1 & -2 \\ & 4/3 & 4/3 \end{array}\right].$$

with $p = [2 \;\; 1 \;\; 3]$ and $q = [2 \;\; 3 \;\; 1 \;\; 4]$. We have

$$Q = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} e_3' \\ e_1' \\ e_2' \\ e_4' \end{bmatrix}$$

so

$$G = Q \begin{bmatrix} L_1 & \\ L_2 & I \end{bmatrix} = Q \left[\begin{array}{cc|cc} 1 & & & \\ -1/3 & 1 & & \\ \hline 0 & 3/4 & 1 & \\ 2/3 & 1/4 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 3/4 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ -1/3 & 1 & 0 & 0 \\ 2/3 & 1/4 & 0 & 1 \end{array}\right].$$

Solve $U_1 V_1 = I_2$ for V_1:

$$\begin{bmatrix} -3 & 1 \\ & 4/3 \end{bmatrix} V_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad V_1 = \begin{bmatrix} -1/3 & 1/4 \\ & 3/4 \end{bmatrix}.$$

Solve $U_1 V_2 + U_2 = 0$ for V_2:

$$\begin{bmatrix} -3 & 1 \\ & 4/3 \end{bmatrix} V_2 = \begin{bmatrix} 2 \\ -4/3 \end{bmatrix}, \qquad V_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

We have

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} e_2' \\ e_1' \\ e_3' \end{bmatrix}$$

so

$$F = P\begin{bmatrix} V_1 & V_2 \\ & I_1 \end{bmatrix} = P\left[\begin{array}{cc|c} -1/3 & 1/4 & -1 \\ & 3/4 & -1 \\ \hline & & 1 \end{array}\right] = \left[\begin{array}{cc|c} 0 & 3/4 & -1 \\ -1/3 & 1/4 & -1 \\ 0 & 0 & 1 \end{array}\right].$$

Problem 7.

Execute the algorithm in the proof of proposition 5, with no pivoting and with complete pivoting, for

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 5 \end{bmatrix}.$$

That is, find two *different* pairs of nonsingular matrices F and G with $AF = GE_\rho$ ($\rho = \rho(A) = 2$).

Characterizations of $\mathcal{R}(A)$ *and* $\mathcal{N}(A)$, *and of* $\mathcal{R}(A')$ *and* $\mathcal{N}(A')$.

We have already shown how to determine if a linear system $Ax = b$ is solvable, exactly, and, if so, how to determine its general solution. The developments which follow will help us to do this in a slightly better way. We essentially need to get "computational handles" on $\mathcal{R}(A)$ and $\mathcal{N}(A)$.

In general, linear systems $Ax = b$ need not be solvable, exactly. Such problems can, in fact, be *much more important* than the exactly solvable ones. In such cases we may ask to determine those $x \in \mathbb{F}^m$ which minimize the sum of squares $\| b - Ax \|^2$. In general, such minimizers x need not be unique, and then there will be solutions with large (gigantic!) norms $\| x \|$. However if we, in addition, minimize $\| x \|^2$ over all solutions of the *least squares problem* $\| b - Ax \|^2 =$ minimum, we shall obtain a unique solution x^*.

In practice, the solutions x to (even nonsingular) linear systems $Ax = b$ can be very sensitive to changes in A, and b. Such problems are "ill-conditioned," or "sick" for short. The following "regularization" procedure can frequently be used to make "sick problems get well." "Approximate" A by a matrix $\widetilde{A}$ with a smaller ρ. Then solve the *generalized least squares problem,*

$$\text{minimize } \|x\|^2 \text{ subject to}$$

$$\|b - \widetilde{A}x\|^2 = \text{minimum}$$

for its shortest solution $\widetilde{x}$. The practical results can be *dramatic.* Frequently, if the "approximation" is done correctly, the "smoothed" solution $\widetilde{x}$ of the "nearby" problem can be much more meaningful, physically, than the solution $x = A^{-1}b$.

These words merely justify our interest in the kinds of problems we are studying. It turns out that this general least squares problem can be solved, as a consequence of the LU theorem, provided we "do for A' as we do for A." Thus we shall "characterize" $\mathcal{R}(A')$ and $\mathcal{N}(A')$ similarly as we "characterize" $\mathcal{R}(A)$ and $\mathcal{N}(A)$.

Let $A \in \mathbb{R}^{n \times m}$ and let

$$AF = GE_\rho : \quad E_\rho = \begin{bmatrix} I_\rho & O \\ O & O \end{bmatrix}$$

with F and G nonsingular. Partition

$$F = \begin{array}{c} \begin{bmatrix} F_1 & F_2 \end{bmatrix} \\ \;\;\rho \quad m-\rho \end{array}, \quad G = \begin{array}{c} \begin{bmatrix} G_1 & G_2 \end{bmatrix} \\ \;\;\rho \quad n-\rho \end{array}$$

$\mathcal{R}(A)$.

Let $y \in \mathcal{R}(A)$. Then $y = Ax$ for some $x \in \mathbb{F}^m$. With the *change of variables*

$$x :=: Fz, \qquad z = \begin{array}{cc} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} & \begin{array}{c} \rho \\ m-\rho \end{array} \end{array}$$

this becomes

$$\begin{aligned} y = Ax = AFz &= GE_\rho z \\ &= G \begin{bmatrix} I_\rho & O \\ O & O \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\ &= \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} z_1 \\ 0 \end{bmatrix} \\ &= G_1 z_1 . \end{aligned}$$

It follows that

$$\mathcal{R}(A) \subseteq \mathcal{R}(G_1) .$$

On the other hand, if $y = G_1 z_1 \in \mathcal{R}(G_1)$ then, working backwards, we see that $y = Ax \in \mathcal{R}(A)$. Thus, in fact,

$$\mathcal{R}(A) = \mathcal{R}(G_1).$$

We also have

$$\mathcal{N}(G_1) = \{0_\rho\} ,$$

since if $G_1 z_1 = 0$, then $G \begin{bmatrix} z_1 \\ 0 \end{bmatrix} = 0$ so $\begin{bmatrix} z_1 \\ 0 \end{bmatrix} = 0$ because G is nonsingular.

$\mathcal{N}(A)$.

If $x \in \mathcal{N}(A)$, then $y = Ax = 0$ above, so $z_1 = 0$. Thus

$$x = \begin{bmatrix} F_1 & F_2 \end{bmatrix} \begin{bmatrix} 0 \\ z_2 \end{bmatrix} = F_2 z_2$$

and so

$$\mathcal{N}(A) \subseteq \mathcal{R}(F_2) .$$

And also vice versa, if $x = F_2 z_2$ then $x = \begin{bmatrix} F_1 & F_2 \end{bmatrix} \begin{bmatrix} 0 \\ z_2 \end{bmatrix} =: Fz$ so

$$\begin{aligned} Ax &= AFz = GE_\rho z \\ &= G \begin{bmatrix} I_\rho & O \\ O & O \end{bmatrix} \begin{bmatrix} 0 \\ z_2 \end{bmatrix} \\ &= G0 \\ &= 0 \end{aligned}$$

and $x \in \mathcal{N}(A)$. Thus

$$\mathcal{N}(A) = \mathcal{R}(F_2) .$$

Also

$$\mathcal{N}(F_2) = \{0_{m-\rho}\} ,$$

since if $F_2 z_2 = 0$ then $F\begin{bmatrix} 0 \\ z_2 \end{bmatrix} = 0$, so $\begin{bmatrix} 0 \\ z_2 \end{bmatrix} = 0$, by the nonsingularity of F.

Now recall how we constructed F and G, from LU factorizations of A, in the proof of proposition 5. We have

$$G := S := QL$$

and

$$F_2 = M := PV$$

with

$$V := \begin{bmatrix} V_2 \\ I \end{bmatrix}: \quad U_1 V_2 + U_2 := O\,.$$

We can do the "dual" construction for A'. For all of this it is *more suggestive* to change to a *notation* which is also "*dual.*" Thus we put

$$K := U'$$

and write the LU factorization as

$$Q'AP = LK' = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} K_1' & K_2' \end{bmatrix}.$$

Then we have the (completely!) "dual" factorization

$$P'A'Q = KL' = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \begin{bmatrix} L_1' & L_2' \end{bmatrix}.$$

We can now work with this factorization to find the "duals" of the matrices $S = G_1$ and $M = F_2$ which we found above for A.

We *summarize*, in the modified notation.

Proposition 6.

Let $0 \neq A \in \mathbb{F}^{n \times m}$ have the LU factorization

$$Q'AP = LK' = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} K_1' & K_2' \end{bmatrix}.$$

Then:

1. $\mathcal{R}(A') = \mathcal{R}(R)$, $\quad \mathcal{N}(R) = \{0_\rho\}$, $\quad$ with $\quad R := PK$.

2. $\mathcal{R}(A) = \mathcal{R}(S)$, $\quad \mathcal{N}(S) = \{0_\rho\}$, $\quad$ with $\quad S := QL$.

3. $\mathcal{N}(A) = \mathcal{R}(M)$, $\quad \mathcal{N}(M) = \{0_{m-\rho}\}$, $\quad$ with $\quad M := PV$

   and

$$V := \begin{bmatrix} V_2 \\ I \end{bmatrix}, \qquad K_1' V_2 + K_2' = O \, .$$

4. $\mathcal{N}(A') = \mathcal{R}(N)$, $\quad \mathcal{N}(N) = \{0_{n-\rho}\}$, $\quad$ with $\quad N := QW$

   and

$$W := \begin{bmatrix} W_2 \\ I \end{bmatrix}, \qquad L_1' W_2 + L_2' = O \, .$$

5. The columns of R are *orthogonal* with the columns of M, and the columns of S are *orthogonal* with the columns of N:

$$R'M = O \quad \text{and} \quad S'N = O \, .$$

□ The *definition* of V shows that $K'V = O$. But $R = PK$ and $M = PV$ so $R'M = K'P'PV = K'V = O$ too. The statement that $S'N = O$ is the "dual" of this. ■

Examples.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix} .$$

1. No pivoting.

$$A = LK' = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} K_1' & K_2' \end{bmatrix} = \begin{bmatrix} 1 & \\ 1 & 1 \\ \hline 1 & -1/3 \\ 1 & 2/3 \end{bmatrix} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ & -3 & -3 \end{array} \right] .$$

a. $\mathcal{R}(A') = \mathcal{R}(R)$, $\quad \mathcal{N}(R) = \{0_2\}$ $\quad$ with $\quad R = K = \begin{bmatrix} 1 & \\ 0 & -3 \\ 1 & -3 \end{bmatrix} .$

b. $\mathcal{R}(A) = \mathcal{R}(S), \qquad \mathcal{N}(S) = \{0_2\} \qquad \text{with} \qquad S = L = \begin{bmatrix} 1 & \\ 1 & 1 \\ 1 & -1/3 \\ 1 & 2/3 \end{bmatrix}.$

c. $\mathcal{N}(A) = \mathcal{R}(M), \qquad \mathcal{N}(M) = \{0_1\} \qquad \text{with} \qquad M = V = \begin{bmatrix} V_2 \\ I \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix},$

as was already known from the previous example!

d. $\mathcal{N}(A') = \mathcal{R}(N), \qquad \mathcal{N}(N) = \{0_2\} \qquad \text{with} \qquad N = W = \begin{bmatrix} W_2 \\ I \end{bmatrix}: \qquad L_1' W_2 + L_2' = O.$

Thus

$$\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} W_2 = \begin{bmatrix} -1 & -1 \\ 1/3 & -2/3 \end{bmatrix},$$

$$W_2 = \begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1/3 & -2/3 \end{bmatrix} = \begin{bmatrix} -4/3 & -1/3 \\ 1/3 & -2/3 \end{bmatrix}$$

and so

$$N = \begin{bmatrix} -4/3 & -1/3 \\ 1/3 & -2/3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Compare this with the matrix N on page 18.

e. *Checks.*

$$R'M = \begin{bmatrix} 1 & 0 & 1 \\ & -3 & -3 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$
S'N = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} -4/3 & -1/3 \\ 1/3 & -2/3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
$$

2. *Complete pivoting.*

$$
Q'AP = LK' = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} K_1' & K_2' \end{bmatrix} = \left[\begin{array}{cc} 1 & \\ -1/3 & 1 \\ \hline 0 & 3/4 \\ 2/3 & 1/4 \end{array}\right] \left[\begin{array}{cc|c} -3 & 1 & -2 \\ & 4/3 & 4/3 \end{array}\right]
$$

with $p = \begin{bmatrix} 2 & 1 & 3 \end{bmatrix}$ and $q = \begin{bmatrix} 2 & 3 & 1 & 4 \end{bmatrix}$, so

$$
P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

a. $\mathcal{R}(A') = \mathcal{R}(\widehat{R}), \quad \mathcal{N}(\widehat{R}) = \{0_2\} \quad \text{with} \quad \widehat{R} = PK = \begin{bmatrix} 1 & 4/3 \\ -3 & 0 \\ -2 & 4/3 \end{bmatrix}.$

b. $\mathcal{R}(A) = \mathcal{R}(\widehat{S}), \quad \mathcal{N}(\widehat{S}) = \{0_2\} \quad \text{with} \quad \widehat{S} = QL = \begin{bmatrix} 0 & 3/4 \\ 1 & 0 \\ -1/3 & 1 \\ 2/3 & 1/4 \end{bmatrix}.$

c. $\mathcal{N}(A) = \mathcal{R}(\widehat{M}), \quad \mathcal{N}(\widehat{M}) = \{0_1\} \quad \text{with} \quad \widehat{M} = PV = P \begin{bmatrix} V_2 \\ I \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix},$

as was already known from the previous example (page 54).

d. $\mathcal{N}(A') = \mathcal{R}(\widehat{N})$, $\quad \mathcal{N}(\widehat{N}) = \{0_2\}$ $\quad$ with $\quad \widehat{N} = QW = Q\begin{bmatrix} W_2 \\ I \end{bmatrix}$: $\quad L_1'W_2 + L_2' = 0.$

Thus

$$\begin{bmatrix} 1 & -1/3 \\ & 1 \end{bmatrix} W_2 = \begin{bmatrix} 0 & -2/3 \\ -3/4 & -1/4 \end{bmatrix},$$

$$W_2 = \begin{bmatrix} 1 & 1/3 \\ & 1 \end{bmatrix}\begin{bmatrix} 0 & -2/3 \\ -3/4 & -1/4 \end{bmatrix} = \begin{bmatrix} -1/4 & -3/4 \\ -3/4 & -1/4 \end{bmatrix}$$

and so

$$\widehat{N} = \begin{bmatrix} 1 & 0 \\ -1/4 & -3/4 \\ -3/4 & -1/4 \\ 0 & 1 \end{bmatrix}.$$

e. *Checks.*

$$\widehat{R}'\widehat{M} = \begin{bmatrix} 1 & -3 & -2 \\ 4/3 & 0 & 4/3 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$\widehat{S}'\widehat{N} = \begin{bmatrix} 0 & 1 & -1/3 & 2/3 \\ 3/4 & 1 & 1 & 1/4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1/4 & -3/4 \\ -3/4 & -1/4 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Problem 8.

a. Find the matrices R, S, M and N for

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 5 \end{bmatrix}$$

using no pivoting, and check that

$$R'M = O, \qquad S'N = O\,.$$

b. Find the corresponding matrices $\widehat{R}$, $\widehat{S}$, $\widehat{M}$ and $\widehat{N}$ for the same A using complete pivoting, and check that

$$\widehat{R}'\widehat{M} = O, \qquad \widehat{S}'\widehat{N} = O\,.$$

Problem 9.

Repeat problem 8 for the matrix A of problem 2 (page 22). The matrices M and $\widehat{M}$ are "empty." Orthogonality relations involving "empty" matrices, here $R'M = O$ and $\widehat{R}'\widehat{M} = O$, should be considered to hold "automatically."

In general, for certain subsets $\mathcal{U}$ of $\mathbb{F}^n$, we have constructed "basis matrices" $B \in \mathbb{F}^{n \times \rho}$ with

$$\mathcal{U} = \mathcal{R}(B), \qquad \mathcal{N}(B) = \{0_\rho\}\,.$$

The following proposition is "soft" but ***very important.*** It gives the most general relation between two "basis matrices" which describe the *same* subset $\mathcal{U}$ of $\mathbb{F}^n$.

The matrix A in the proposition is ***not the same matrix*** as the A we have been considering in order to "solve" $Ax = b$. Here is a table which allows a translation from the matrices we have been considering to the notation of the proposition, which is along the top row:

$\mathcal{U}$	B	A	n	ρ
$\mathcal{R}(A')$	R	$\widehat{R}$	m	ρ
$\mathcal{R}(A)$	S	$\widehat{S}$	n	ρ
$\mathcal{N}(A)$	M	$\widehat{M}$	m	$m-\rho$
$\mathcal{N}(A')$	N	$\widehat{N}$	n	$n-\rho$

Key Proposition 7 ("change of basis").

Let $A \in \mathbb{F}^{n \times \rho}$ and $B \in \mathbb{F}^{n \times \sigma}$. Then:

1. If $A = BC$ with C nonsingular then $\rho = \sigma$ and

$$\mathcal{R}(A) = \mathcal{R}(B), \qquad C\mathcal{N}(A) = \mathcal{N}(B).$$

In particular,

$$\mathcal{N}(A) = \{0_\rho\} \;\Leftrightarrow\; \mathcal{N}(B) = \{0_\rho\}\,.$$

2. If $\mathcal{R}(A) = \mathcal{R}(B)$, $\mathcal{N}(A) = \{0_\rho\}$ and $\mathcal{N}(B) = \{0_\sigma\}$, then $\rho = \sigma$ and there's a unique nonsingular matrix C with

$$A = BC\ .$$

□

1.

a. Since $A = BC$ then $C \in \mathbb{F}^{\sigma \times \rho}$. Since C is nonsingular then $\rho = \sigma$.

b. We have

$$\begin{aligned}\mathcal{R}(A) &= \{Ax:\ x \in \mathbb{F}^{\rho}\} \\ &= \{BCx:\ x \in \mathbb{F}^{\rho}\}\ .\end{aligned}$$

Make the *change of variables*

$$z :=: Cx\ .$$

Since C is nonsingular, as x varies over *all* of $\mathbb{F}^{\rho}$ so does z, and vice versa. Thus

$$\mathcal{R}(A) = \{Bz:\ z \in \mathbb{F}^{\rho}\} = \mathcal{R}(B).$$

c. With the same change of variables we have

$$\begin{aligned}\mathcal{N}(B) &= \{z \in \mathbb{F}^{\rho}:\ Bz = 0_n\} \\ &= \{Cx \in \mathbb{F}^{\rho}:\ BCx = 0_n\} \\ &= C\{x \in \mathbb{F}^{\rho}:\ Ax = 0_n\} \\ &= C\mathcal{N}(A)\ .\end{aligned}$$

d. We have

$$\mathcal{N}(A) = \{0_\rho\} \ \Rightarrow\ \mathcal{N}(B) = C\{0_\rho\} = \{0_\rho\}\ .$$

On the other hand

$$\begin{aligned}\mathcal{N}(B) = \{0_\rho\} \ &\Rightarrow\ C\mathcal{N}(A) = \{0_\rho\} \\ &\Rightarrow\ \mathcal{N}(A) = \{0_\rho\}\end{aligned}$$

since C is nonsingular.

2. First suppose that $\rho \geq \sigma$.

a. Let $C \in \mathbb{F}^{\sigma \times \rho}$ solve $BC = A$. Thus the columns of C solve

$$Bc_i = a_i\,, \qquad i = 1, 2, \ldots, \rho.$$

Since

$$a_i = Ae_i \in \mathcal{R}(A) = \mathcal{R}(B)\ ,$$

all these linear systems are solvable, and uniquely so since $\mathcal{N}(B) = \{0_\sigma\}$.

b. We have $\mathcal{N}(C) = \{0_\rho\}$. For if $Cx = 0$ then $Ax = BCx = B0 = 0$ and $x = 0$ since $\mathcal{N}(A) = \{0_\rho\}$. Since $\mathcal{N}(C) = \{0_\rho\}$ then $\rho \le \sigma$ (key fact GF2, page 10). But we assumed that $\rho \ge \sigma$. Thus $\rho = \sigma$. C is square, and $\mathcal{N}(C) = \{0_\rho\}$ means that C is nonsingular.

c. Suppose now that $\rho < \sigma$. Interchange the roles of A and B above to find a matrix $D \in \mathbb{F}^{\rho \times \sigma}$ with $B = AD$ and $\mathcal{N}(D) = \{0_\sigma\}$. Then $\sigma \le \rho$, a contradiction. Thus this case cannot occur. ∎

Examples.

We exhibit the "change of basis" matrices C for the matrices A and B of the table preceding the statement of proposition 7. We use the pairs of "basis matrices" from the previous example.

a. $\mathcal{R}(A')$.

$$R = \begin{bmatrix} 1 & \\ 0 & -3 \\ 1 & -3 \end{bmatrix}, \qquad \widehat{R} = \begin{bmatrix} 1 & 4/3 \\ -3 & 0 \\ -2 & 4/3 \end{bmatrix} \stackrel{?}{=} RC.$$

Equate the first two rows to get

$$\begin{bmatrix} 1 & \\ 0 & -3 \end{bmatrix} C = \begin{bmatrix} 1 & 4/3 \\ -3 & 0 \end{bmatrix}.$$

Thus

$$C = \begin{bmatrix} 1 & \\ & -1/3 \end{bmatrix} \begin{bmatrix} 1 & 4/3 \\ -3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 4/3 \\ 1 & 0 \end{bmatrix}.$$

Check:

$$RC = \begin{bmatrix} 1 & \\ 0 & -3 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 4/3 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 4/3 \\ -3 & 0 \\ -2 & 4/3 \end{bmatrix} = \widehat{R}.$$

b. $\mathcal{R}(A)$.

$$S = \begin{bmatrix} 1 & \\ 1 & 1 \\ 1 & -1/3 \\ 1 & 2/3 \end{bmatrix}, \qquad \widehat{S} = \begin{bmatrix} 0 & 3/4 \\ 1 & 0 \\ -1/3 & 1 \\ 2/3 & 1/4 \end{bmatrix} \stackrel{?}{=} SC.$$

Equate the first two rows:

$$\begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix} C = \begin{bmatrix} 0 & 3/4 \\ 1 & 0 \end{bmatrix}.$$

Thus

$$C = \begin{bmatrix} 1 & \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 3/4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 3/4 \\ 1 & -3/4 \end{bmatrix}.$$

Check:

$$SC = \begin{bmatrix} 1 & \\ 1 & 1 \\ 1 & -1/3 \\ 1 & 2/3 \end{bmatrix} \begin{bmatrix} 0 & 3/4 \\ 1 & -3/4 \end{bmatrix} = \begin{bmatrix} 0 & 3/4 \\ 1 & 0 \\ -1/3 & 1 \\ 2/3 & 1/4 \end{bmatrix} = \widehat{S}.$$

c. $\mathcal{N}(A)$.

$$M = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \quad \widehat{M} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = MC$$

with

$$C = \begin{bmatrix} 1 \end{bmatrix}$$

(an "accident").

d. $\mathcal{N}(A')$.

$$N = \begin{bmatrix} -4/3 & -1/3 \\ 1/3 & -2/3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \widehat{N} = \begin{bmatrix} 1 & 0 \\ -1/4 & -3/4 \\ -3/4 & -1/4 \\ 0 & 1 \end{bmatrix} \stackrel{?}{=} NC.$$

Equate the last two rows:

$$C = \begin{bmatrix} -3/4 & -1/4 \\ 0 & 1 \end{bmatrix}.$$

Check:

$$NC = \begin{bmatrix} -4/3 & -1/3 \\ 1/3 & -2/3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -3/4 & -1/4 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1/4 & -3/4 \\ -3/4 & -1/4 \\ 0 & 1 \end{bmatrix} = \widehat{N}$$

Problem 10.

Find the "change of basis" matrices C for the "basis matrix" pairs you found in problem 8, similarly.

Problem 11.

Same problem, for the pairs of problem 9, except for M and $\widehat{M}$.

The Language of Linear Algebra (LAL).

Let $A \in \mathbb{F}^{n \times m}$ and $b \in \mathbb{F}^n$. We study the *geometry* of "*solving*" the linear system $Ax = b$ for $x \in \mathbb{F}^m$.

Basic definitions and consequences.

1. *Matrix-vector multiplication.*

$$Ax = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{bmatrix}$$

$$= a_1\xi_1 + a_2\xi_2 + \dots + a_m\xi_m$$

is a linear combination (ℓc) *of the columns of* A.

2. (Not so) roughly speaking, a *linear space* (or *vector space*) is an algebraic system in which it "makes sense" to form linear combinations. We refrain from giving the rather long list of axioms for a general linear space as we shall mostly work with the primary, yet typical, example of the linear space $\mathbb{F}^n$. Linear combinations are broken down into the two more basic algebraic operations of *scalar multiplication* and *vector addition*. The scalars must come from a "field" $\mathbb{F}$. We shall usually have $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, but the cases $\mathbb{F} = \mathbb{Q}$, the *rational numbers*, and $\mathbb{F} = GF(p)$, the *finite field* of integers modulo p, a prime number, are also useful in applications. In our treatment we shall take $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, but all results carry over directly to more general "fields."

The $n \times m$ matrices $\mathbb{F}^{n \times m}$ also form a linear space with respect to the scalar "field" $\mathbb{F}$, but this linear space is "algebraically the same as," or "isomorphic with" the linear space $\mathbb{F}^{mn}$ under the "*stacking*" correspondence

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} \leftrightarrow \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}.$$

That is, we may form ℓcs of matrices by forming the corresponding ℓcs of the associated "stacked" vectors. In fact, A is stored in the computer as **a**! In matlab: $A(:) = \mathbf{a}$.

Other kinds of linear spaces may consist of linear spaces *of functions.* For instance, the set of twice continuously differentiable solutions of the linear *homogeneous* second-order ODE

$$\alpha\xi'' + \beta\xi' + \gamma\xi = 0, \qquad 0 \le \tau \le 1\,,$$

forms a linear space. Here α, β and γ are given functions defined for $0 \le \tau \le 1$. We shall not emphasize such types of linear spaces here, although our results carry over to them as well.

3. A sub*set* $\mathcal{U}$ of a linear space $\mathcal{V}$ is a ***subspace*** of $\mathcal{V}$ if $\mathcal{U}$ is a linear space in its own right. For this it is necessary and sufficient that $\mathcal{U}$ be *closed* under the formation of ℓcs:

$$\boxed{\begin{array}{c} \alpha_1, \alpha_2 \in \mathbb{F}, \qquad x_1, x_2 \in \mathcal{U} \\ \Rightarrow\ \alpha_1 x_1 + \alpha_2 x_2 \in \mathcal{U} \end{array}}\ .$$

Examples.

a. $\mathcal{V} = \mathbb{R}^3$ has *four kinds* of subspaces $\mathcal{U}$:

0) $\{0_3\}$,

1) lines through 0_3,

2) planes containing 0_3,

3) $\mathbb{R}^3$.

b. The lower triangular matrices $\mathbb{F}_\ell^{n \times n}$ form a subspace of $\mathbb{F}^{n \times n}$. The upper triangular matrices $\mathbb{F}_u^{n \times n}$ form a subspace of $\mathbb{F}^{n \times n}$. The diagonal matrices $\mathbb{F}_d^{n \times n}$ form a subspace of $\mathbb{F}^{n \times n}$. $\mathbb{F}_d^{n \times n}$ is also a subspace of *each* of the subspaces $\mathbb{F}_\ell^{n \times n}$ and $\mathbb{F}_u^{n \times n}$. In fact, $\mathbb{F}_d^{n \times n}$ is the *intersection*

$$\mathbb{F}_d^{n \times n} = \mathbb{F}_\ell^{n \times n} \cap \mathbb{F}_u^{n \times n}\,.$$

c. The real symmetric matrices $(A = A')$ $\mathbb{R}_s^{n \times n}$ form a subspace of $\mathbb{R}^{n \times n}$. The real skew symmetric matrices $(A + A' = O)$ $\mathbb{R}_k^{n \times n}$ form a subspace of $\mathbb{R}^{n \times n}$. The intersection of these two subspaces is

$$\mathbb{R}_s^{n \times n} \cap \mathbb{R}_k^{n \times n} = \{0_{n \times n}\},$$

i.e., the zero matrix of $\mathbb{R}^{n \times n}$.

d. If $\mathcal{U}$ and $\mathcal{V}$ are subspaces of a linear space W, then the *intersection*

$$\mathcal{U} \cap \mathcal{V} := \{t:\ t \in \mathcal{U} \quad \textit{and} \quad t \in \mathcal{V}\}$$

and *sum*

$$\mathcal{U} + \mathcal{V} := \{u + v:\ u \in \mathcal{U},\ v \in \mathcal{V}\}$$

are subspaces of $\mathcal{W}$, but the *union*

$$\mathcal{U} \cup \mathcal{V} := \{t:\ t \in \mathcal{U} \quad \textit{or} \quad t \in \mathcal{V}\}$$

is *not*, in general, a subspace of $\mathcal{W}$. In fact, the sum $\mathcal{U} + \mathcal{V}$ is the *smallest* subspace of $\mathcal{W}$ which contains the union $\mathcal{U} \cup \mathcal{V}$.

e. Let w be a fixed vector in $\mathbb{F}^n$. Then

$$w^{\perp} := \left\{x \in \mathbb{F}^n:\ w'x = 0\right\}$$

is a subspace of $\mathbb{F}^n$. Let $\mathcal{W}$ be a sub*set* of $\mathbb{F}^n$. Then

$$w^{\perp} := \left\{x \in \mathbb{F}^n:\ w'x = 0 \text{ for } \textit{all } w \in \mathcal{W}\right\}$$

is a subspace of $\mathbb{F}^n$. In fact we have

$$\mathcal{W}^{\perp} = \cap \left\{w^{\perp}:\ w \in \mathcal{W}\right\}.$$

f. One way to state the *Fundamental Theorem of Linear Algebra* (*FTLA*) is that $\mathcal{N}(A)$ and $\mathcal{R}(A')$ are subspaces of $\mathbb{F}^m$ with

$$\mathcal{R}(A') = \mathcal{N}(A)^{\perp}$$

and, "dually," $\mathcal{N}(A')$ and $\mathcal{R}(A)$ are subspaces of $\mathbb{F}^n$ with

$$\mathcal{N}(A') = \mathcal{R}(A)^{\perp}.$$

We'll establish this result *constructively*, that is with an *algorithm* based on our work above.

Problem 12.

Prove property d, that $\mathcal{U} \cap \mathcal{V}$ and $\mathcal{U} + \mathcal{V}$ are subspaces, but that $\mathcal{U} \cup \mathcal{V}$ is *not*, in general, a subspace. Also prove the first part of property e, that $w^{\perp}$ is a subspace.

4. The mapping

$$A\colon \mathbb{F}^m \to \mathbb{F}^n$$

$$x \to Ax$$

is a *linear transformation*, that is

$$A(\alpha_1 x_1 + \alpha_2 x_2) \equiv \alpha_1 A x_1 + \alpha_2 A x_2$$

(*identically*, for *all* scalars α_1, $\alpha_2 \in \mathbb{F}$ and vectors x_1, $x_2 \in \mathbb{F}^m$). Note that this can also be written as

$$A\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \equiv \begin{bmatrix} Ax_1 & Ax_2 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.$$

Thus, this "*linearity of* A" is just the statement that

$$A\begin{bmatrix} x_1 & x_2 \end{bmatrix} \equiv \begin{bmatrix} Ax_1 & Ax_2 \end{bmatrix}.$$

All of the results of this section are "soft," or conceptual. They follow directly from the definitions and from properties which we have already established. Here is our first "soft" result.

Proposition 8.

1. $\mathcal{R}(A)$ is a subspace of $\mathbb{F}^n$. The linear system $Ax = b$ is solvable if and only if $b \in \mathcal{R}(A)$.
2. $\mathcal{N}(A)$ is a subspace of $\mathbb{F}^m$. If $Ax = b$ is solvable, by the particular solution x_p, then the set of *all* solutions of $Ax = b$ is

$$x_p + \mathcal{N}(A) := \{x_p + z\colon\ Az = 0\}\ .$$

□

1. We show that $\mathcal{R}(A)$ is *closed* under the formation of ℓcs. Let α_1, $\alpha_2 \in \mathbb{F}$ and y_1, $y_2 \in \mathcal{R}(A)$. Then $y_k = Ax_k$ with $x_k \in \mathbb{F}^m$ $(k = 1, 2)$. *By the linearity of* A,

$$\begin{aligned} \alpha_1 y_1 + \alpha_2 y_2 &= \alpha_1 A x_1 + \alpha_2 A x_2 \\ &= A(\alpha_1 x_1 + \alpha_2 x_2) \\ &\in \mathcal{R}(A). \end{aligned}$$

$\mathcal{R}(A)$ is closed under the formation of ℓcs and is thus a subspace of $\mathbb{F}^n$. The second statement of #1 is just the *definition* of $\mathcal{R}(A)$.

2. We first show that $\mathcal{N}(A)$ is closed under the formation of ℓcs. Let $\alpha_1, \alpha_2 \in \mathbb{F}$ and $z_1, z_2 \in \mathcal{N}(A)$. Thus $Az_1 = Az_2 = 0_n$. *By the linearity of* A,

$$\begin{aligned} A(\alpha_1 z_1 + \alpha_2 z_2) &= \alpha_1 A z_1 + \alpha_2 A z_2 \\ &= \alpha_1 0_n + \alpha_2 0_n \\ &= 0_n . \end{aligned}$$

Thus $\alpha_1 z_1 + \alpha_2 z_2 \in \mathcal{N}(A)$, $\mathcal{N}(A)$ is closed under the formation of ℓcs, and is thus a subspace of $\mathbb{F}^m$.

Now suppose $Ax = b$ is solvable, with $Ax_p = b$. If $z \in \mathcal{N}(A)$, that is $Az = 0$, then also $A(x_p + z) = Ax_p + Az = b + 0_n = b$, by the linearity of A. Thus, every vector $x_p + z$ of $x_p + \mathcal{N}(A)$ solves $Ax = b$. Vice versa, suppose $Ax = b$. Define $z := x - x_p$. Then, by the *linearity* of A, $Az = A(x - x_p) = Ax - Ax_p = b - b = 0_n$, so $z \in \mathcal{N}(A)$ and $x = x_p + z \in x_p + \mathcal{N}(A)$. Thus every solution of $Ax = b$ is a member of $x_p + \mathcal{N}(A)$. ■

3. We have

$$\mathcal{R}(A) = \{Ax:\ x \in \mathbb{F}^m\}$$

in *matrix language*. Now

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}$$

and

$$Ax = a_1\xi_1 + a_2\xi_2 + \dots + a_m\xi_m$$

is a linear combination of the columns of A. For this reason $\mathcal{R}(A)$ is frequently called the *column space* of A. The columns of A are vectors in the linear space $\mathbb{F}^n$. The Linear Algebra Language (LAL) is that the columns of A *span* the subspace $\mathcal{U} = \mathcal{R}(A)$ of $\mathcal{V} = \mathbb{F}^n$.

Only superficially more generally, if $v_1, v_2, \dots, v_m$ are vectors from a linear space $\mathcal{V}$, then

$$\mathcal{U} := \{\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m:\ \alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{F}\}$$

is a subspace of $\mathcal{V}$ and the vectors $v_1, v_2, \dots, v_m$ are said to *span* $\mathcal{U}$. We write this as

$$\mathcal{U} =: \operatorname{span}\{v_1, v_2, \dots, v_m\} .$$

Thus, in our (concrete) case we have

$$\mathcal{R}(A) = \operatorname{span}\{a_1, a_2, \dots, a_m\}$$

with the matrix notation the simplest.

4. In matrix language,

$$\mathcal{N}(A) = \{x \in F^m: \; Ax = 0_n\}$$

is the set of all solutions of the homogeneous system. We have the unique solution $x = 0_m$ if and only if

$$\mathcal{N}(A) = \{0_m\} .$$

The LAL for this is that the columns of A are *linearly independent* (ℓi). In other words, if

$$a_1\xi_1 + a_2\xi_2 + \ldots + a_m\xi_m = 0$$

then

$$\xi_1 = \xi_2 = \ldots = \xi_m = 0 .$$

Otherwise, when $\mathcal{N}(A) \neq \{0_m\}$, the columns of A are *linearly dependent* (ℓd). This means that there is an $x \neq 0$ so that $Ax = 0$. In LAL, there are scalars $\xi_1, \xi_2, \ldots, \xi_m$, not all zero, so that

$$a_1\xi_1 + a_2\xi_2 + \ldots + a_m\xi_m = 0 .$$

Then, one of the columns a_i can be written as an ℓc of the *remaining* columns.

Any subset of a (finite) ℓi set of vectors in F^n is also an ℓi set. For if $Ax = 0 \Rightarrow x = 0$, and if we partition $A = [A_1 \;\; A_2]$, then $A_1x_1 = 0$ implies

$$0 = A_1x_1 = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

from which we obtain

$$\begin{bmatrix} x_1 \\ 0 \end{bmatrix} = 0, \text{ i.e., } x_1 = 0.$$

By logical convention the *empty set* of vectors in F^n is also an ℓi set.

The concepts of "ℓi" and "ℓd" extend, again with superficial generality, to general linear spaces. The subset $\{v_1, v_2, v_3, \ldots\}$ of vectors from $\mathcal{V}$ is an ℓi set if, whenever any *finite* ℓc of the v_i is the zero vector, then *all* the coefficients in the ℓc are zeros. And ℓd is the *opposite* of ℓi.

5. *Basis.*

The columns of B form a *basis* for the subspace $\mathcal{U}$ of $\mathcal{V} = F^n$ if

$$\mathcal{U} = \mathcal{R}(B) \quad \textit{and} \quad \mathcal{N}(B) = \{0\} .$$

Thus, the columns of B form a *linearly independent spanning set for* $\mathcal{U}$.

There is a conceptually easy, but very inefficient, way to obtain a basis for $\mathcal{U} = \mathcal{R}(A)$ when $A \in \mathbb{F}^{n \times m}$ is an arbitrary matrix. If $\mathcal{N}(A) = \{0\}$ we are done, with $B := A$. If not then there is a vector $x \neq 0$ with $Ax = 0$. Suppose, by reordering the columns, if necessary, that $\xi_m \neq 0$. In fact, by the homogeneity, we may take $\xi_m := -1$. Then a_m is an ℓc of $a_1, a_2, \ldots, a_{m-1}$ and thus so is every $y \in \mathcal{U} = \mathcal{R}(A)$. Hence $\mathcal{U} = \mathcal{R}(A_{m-1})$ with $A_{m-1} := [a_1 \ \ a_2 \ \ \ldots \ \ a_{m-1}]$. If $\mathcal{N}(A_{m-1}) = \{0_{m-1}\}$, we are done, with $B := A_{m-1}$. Otherwise we repeat the process. We eventually arrive at some $B := A_k = [a_1 \ \ a_2 \ \ \ldots \ \ a_k]$ $(k \leq m)$ for which

$$\mathcal{U} = \mathcal{R}(B), \qquad \mathcal{N}(B) = \{0\} .$$

If $A = O$ then $k = 0$, $B = \emptyset$.)

Note that the basis for $\mathcal{R}(A)$ obtained here is a *subset* of the original spanning set for $\mathcal{R}(A)$, the column of A. This property is sometimes desired.

The argument shows that *a basis for* $\mathcal{U}$ *is a minimal spanning set for* $\mathcal{U}$, that is it contains a minimal *number* of vectors.

The same definition of basis works in general linear spaces $\mathcal{V}$. The set $\{v_1, v_2, v_3, \ldots\}$ is a *basis* for $\mathcal{V}$ if it is a linearly independent spanning set for $\mathcal{V}$. A linear space $\mathcal{V}$ is *finite dimensional* if it is spanned by a finite set of vectors. The above construction then also works to eliminate "redundant" vectors from the spanning set to obtain a *finite* basis for $\mathcal{V}$.

The first part of the following "soft" proposition shows that a basis for a subspace $\mathcal{U}$ of $\mathbb{F}^n$ is also a *maximal linearly independent set* of vectors from $\mathcal{U}$. The second part is an easy consequence of proposition 7, which is much more explicit.

Proposition 9.

Let $\mathcal{U} \neq \{0_n\}$ be a subspace of $\mathbb{F}^n$. Then:

1. Every linearly independent set of vectors in $\mathcal{U}$, including the empty set, can be extended to form a basis for $\mathcal{U}$.

2. Every basis for $\mathcal{U}$ contains the same number of vectors. This number is called the *dimension* of $\mathcal{U}$ and is denoted by dim $\mathcal{U}$.

□

1. Let $b_1, b_2, \ldots, b_k$ be the given ℓi set $(k \geq 0)$ and put $B_k := [b_1 \ \ b_2 \ \ \ldots \ \ b_k]$. We have $\mathcal{N}(B_k) = \{0\}$ and $\mathcal{R}(B_k) \subseteq \mathcal{U}$. If $\mathcal{R}(B_k) = \mathcal{U}$ we are done: the columns of $B := B_k$ form a basis for $\mathcal{U}$. Otherwise there is a vector $b_{k+1} \in \mathcal{U} \setminus \mathcal{R}(B_k)$, i.e., in $\mathcal{U}$ but *not* in $\mathcal{R}(B_k)$. Let

$B_{k+1} := [B_k \;\; b_{k+1}]$. We show that $\mathcal{N}(B_{k+1}) = \{0\}$. Suppose $\begin{bmatrix} B_k & b_{k+1} \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix} = 0$, that is

$B_k x + b_{k+1}\xi = 0$. If $\xi \neq 0$, then $b_{k+1} = -B_k x/\xi \in \mathcal{R}(B_k)$, contrary to the choice of b_{k+1}. Thus $\xi = 0$, $B_k x = 0$, so $x = 0$ since $\mathcal{N}(B_k) = \{0\}$. Thus $\mathcal{N}(B_{k+1}) = \{0\}$, as required. If $\mathcal{R}(B_{k+1}) = \mathcal{U}$ we are done, as above. We continue, repeating the process. It must stop with some $B = B_\rho$ for which $\mathcal{N}(B) = \{0\}$ and $\mathcal{R}(B) = \mathcal{U}$, since any linear independent set of vectors in $\mathbb{F}^n$ contains at most n vectors (key fact GF2, page 10). In particular, *every* subspace $\mathcal{U}$ of $\mathbb{F}^n$ has a basis!

■

Examples.

a. The *axis vectors* $e_1, e_2, \ldots, e_n$ in $\mathbb{R}^n$ form a basis, the *standard basis*, for $\mathbb{R}^n$, and even for $\mathbb{C}^n$. Thus, $\dim \mathbb{R}^n = \dim \mathbb{C}^n = n$.

b. The mn *outer products* $e_j e_i'$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, where the e_i are in $\mathbb{R}^m$ and the e_j are in $\mathbb{R}^n$, form a basis for the linear space $\mathbb{R}^{n\times m}$, and even for $\mathbb{C}^{n\times m}$. Thus, $\dim \mathbb{R}^{n\times m} = \dim \mathbb{C}^{n\times m} = mn$.

c. Consider the real symmetric matrices $\mathbb{R}_s^{n\times n}$. For example if $H \in \mathbb{R}_s^{3\times 3}$ then

$$H = \begin{bmatrix} \eta_{11} & \eta_{21} & \eta_{31} \\ \eta_{21} & \eta_{22} & \eta_{32} \\ \eta_{31} & \eta_{32} & \eta_{33} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\eta_{11} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\eta_{22} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\eta_{33}$$

$$+ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\eta_{21} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\eta_{31} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\eta_{32}.$$

Moreover, the $\frac{n(n+1)}{2} = 6$ matrices given here are clearly ℓi. Only slightly more generally we see that

$$\dim \mathbb{R}_s^{n\times n} = \frac{n(n+1)}{2}.$$

In the same way we see that

$$\dim \mathbb{R}_k^{n\times n} = \frac{n(n-1)}{2},$$

since the diagonal elements of $K \in \mathbb{R}_k^{n\times n}$ are all zero.

d. Likewise we have

$$\dim \mathbb{F}_\ell^{n\times n} = \frac{n(n+1)}{2} = \dim \mathbb{F}_u^{n\times n}$$

and

$$\dim \mathbb{F}_d^{n\times n} = n\ .$$

e. Proposition 6 shows that the *four fundamental subspaces*, $\mathcal{R}(A)$, $\mathcal{N}(A)$, $\mathcal{R}(A')$ and $\mathcal{N}(A')$, associated with a matrix $A \in \mathbb{F}^{n\times m}$ have bases with ρ, $m-\rho$, ρ and $n-\rho$ vectors, respectively. These bases are the columns of the matrices which we called S, M, R and N, respectively. Thus we have

$$\dim \mathcal{R}(A) = \rho = \dim \mathcal{R}(A')$$

and

$$\dim \mathcal{N}(A) = m-\rho, \qquad \dim \mathcal{N}(A') = n-\rho\ .$$

Also, if $Ax = b$ is solvable, by the particular solution $x_p \in \mathbb{F}^m$, then the set of *all* solutions of $Ax = b$ is

$$x_p + \mathcal{N}(A) = \{x_p + Nz:\ z \in \mathbb{F}^{m-\rho}\}\ .$$

Definitions.

$$\begin{aligned} \operatorname{rank} A &:= \textit{rank} \text{ of } A \\ &:= \dim \mathcal{R}(A) \end{aligned}$$

and

$$\begin{aligned} \operatorname{null} A &:= \textit{nullity} \text{ of } A \\ &:= \dim \mathcal{N}(A). \end{aligned}$$

Proposition 10. For $A \in \mathbb{F}^{n\times m}$,

$$\operatorname{rank} A = \operatorname{rank} A'$$

and

$$\operatorname{rank} A + \operatorname{null} A = m \qquad \text{(Sylvester's law of nullity).}$$

□ Trivial!

■

Note that Sylvester's law, as formally stated, relates the dimensions of the two subspaces, $\mathcal{R}(A)$ and $\mathcal{N}(A)$, of the two *different* linear spaces, $\mathbb{F}^n$ and $\mathbb{F}^m$, respectively. However, because of the first relation, rank A = rank A′, we also have

$$\dim \mathcal{R}(A') + \dim \mathcal{N}(A) = m$$

as well as its "dual,"

$$\dim \mathcal{R}(A) + \dim \mathcal{N}(A') = n.$$

Thus the *sum* of the dimensions of the subspaces $\mathcal{R}(A)$ and $\mathcal{N}(A')$, *both* subspaces of $\mathbb{F}^n$, is n, the dimensions of $\mathbb{F}^n$ itself. How can this be? Might the bases for $\mathcal{R}(A)$ and $\mathcal{N}(A)$, *together*, form a basis for *all* of $\mathbb{F}^n$? We shall show that the answer is "yes." This is one form of the FTLA.

We give another proposition, actually some basic facts which are of interest in their own right. Part 4, $\mathcal{R}(A) \perp \mathcal{N}(A')$ says that *every* vector $y \in \mathcal{R}(A)$ is orthogonal with *every* vector $v \in \mathcal{N}(A')$.

Proposition 11.

1. Let $\mathcal{U}$ and $\mathcal{V}$ be subspaces of $\mathbb{F}^n$, with $\mathcal{U}$ a subspace of $\mathcal{V}$. Then,

$$\dim \mathcal{U} \subseteq \dim \mathcal{V}\ ,$$

with equality if and only if

$$\mathcal{U} = \mathcal{V}.$$

2. rank AB $\leq$ min {rank A, rank B}

3. $$\mathcal{R}(A'A) = \mathcal{R}(A')\ .$$

Thus the *normal equations*

$$A'Ax = A'b$$

are always solvable.

4. $$\mathcal{R}(A) \perp \mathcal{N}(A').$$

□

1. Extend a basis for $\mathcal{U}$ to a basis for $\mathcal{V}$, as in the proof of proposition 9.1 (there we have $\mathcal{V} = \mathbb{F}^n$). We have $\mathcal{U} = \mathcal{V}$ exactly when no extension is necessary.

2. We have rank AB = dim $\mathcal{R}(AB)$. Moreover, if $B \in \mathbb{F}^{m \times k}$ say, then $\mathcal{R}(AB) = \{ABx:\ x \in \mathbb{F}^n\} \subseteq \mathcal{R}(A)$. That is, $\mathcal{R}(AB)$ is a subspace of $\mathcal{R}(A)$. By #1, rank AB $\leq$ rank A. But, how do we obtain the *other* inequality, rank AB $\leq$ rank B? Answer:

transpose the matrices and use the first statement of proposition 10:

rank $AB = \text{rank}(AB)' = \text{rank } B'A' \leq \text{rank } B' = \text{rank } B$, by what was shown above $(A \leftrightarrow B')$.

3. $\mathcal{R}(A'A)$ is a subspace of $\mathcal{R}(A')$. By 1, and the definition of rank, we have $\mathcal{R}(A'A) = \mathcal{R}(A')$ if and only if rank $A'A$ = rank A' (= rank A). By Sylvester's law, for $A \in \mathbb{F}^{n \times m}$,

$$\text{rank } A'A + \text{null } A'A = m$$

and

$$\text{rank } A + \text{null } A = m .$$

Thus, our result will be proved if we can show that

$$\text{null } A'A = \text{null } A.$$

Now, in general, $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$ $(Bx = 0 \Rightarrow ABx = 0)$. Thus, by 1, our statement is equivalent with the statement that

$$\mathcal{N}(A'A) = \mathcal{N}(A)$$

and, as just above, we know that $\mathcal{N}(A) \subseteq \mathcal{N}(A'A)$. But if $A'Ax = 0$ then $\| Ax \|^2 = x'A'Ax = x'0 = 0$, so $Ax = 0$ and $\mathcal{N}(A'A) \subseteq \eta(A)$. [This argument doesn't work with general "fields," only with "subfields" of $\mathbb{C}$, e.g., $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Q}$.]

4. This *important* result is *easy*. If $y \in \mathcal{R}(A)$ and $v \in \mathcal{N}(A')$ then $y = Ax$, for some $x \in \mathbb{F}^m$, and $A'v = 0$. Thus, by the reverse order rule,

$$y'v = (Ax)'v = x'A'v = x'0 = 0,$$

so y and v are *orthogonal.*

■

The Fundamental Theorem of Linear Algebra.

Let $A \in \mathbb{F}^{n \times m}$. Then:

1. Every $b \in \mathbb{F}^n$ has a unique decomposition

$$b = p + r, \qquad p \in \mathcal{R}(A), \qquad r \in \mathcal{N}(A') .$$

2. p is the unique solution of the least squares problem

$$\min\{ \| b - y \|^2 : \ y \in \mathcal{R}(A) \} ,$$

with minimum value

$$\| b - p \|^2 = \| r \|^2 .$$

3. Every $x_p \in \mathbb{F}^m$ has a unique decomposition

$$x_p = x^* + z : \; x^* \in \mathcal{R}(A'), \qquad z \in \mathcal{N}(A) .$$

4. If x_p solves $Ax = p$ then x^* is the unique solution of

$$\min\{ \|x\|^2 : \; Ax = p\}$$

(with minimum value $\|x^*\|^2$).

□

We use the basis matrices of proposition 6.

1. Since the columns of S form a basis for $\mathcal{R}(A)$, every $p \in \mathcal{R}(A)$ can be written uniquely as

$$p = Su, \qquad u \in \mathbb{F}^{\rho} .$$

Since the columns of N form a basis for $\mathcal{N}(A')$, every $r \in \mathcal{N}(A')$ can be written uniquely as

$$r = Nv, \qquad v \in \mathbb{F}^{n-\rho} .$$

Thus the equation $p + r = b$ is equivalent with the linear system

$$Su + Nv = b$$

of n equations in n unknowns. Moreover, since $S = QL$ and $N = QW$ with Q a permutation matrix, this system is equivalent with

$$Lu + Wv = Q'b ,$$

that is with

$$\begin{bmatrix} L_1 & W_2 \\ L_2 & I \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = Q'b . \qquad (*)$$

We show that

$$\begin{bmatrix} L & W \end{bmatrix} = \begin{bmatrix} L_1 & W_2 \\ L_2 & I \end{bmatrix}$$

is nonsingular. This is surely true when $\rho = n$, for then $L_2 = \emptyset$, $W_2 = \emptyset$, and the matrix is $L_1 =$ [illegible] . If $\rho < n$ then

$$W = \begin{bmatrix} W_2 \\ I \end{bmatrix} : \qquad L_1' W_2 + L_2' = O ,$$

so

$$L'W = O, \qquad W'L = O\,.$$

Consider the homogeneous system

$$Lu + Wv = 0\,.$$

Premultiplication by L' and W', respectively, give

$$L'Lu = 0, \qquad W'Wv = 0\,.$$

But $L'L$ and $W'W$ are positive definite so $u = 0_\rho$, $v = 0_{n-\rho}$. So $\mathcal{N}([L \;\; W]) = \{0_n\}$ and $[L \;\; W]$ is nonsingular. Thus u and v solving (*) are uniquely determined and then so are $p = Su$ and $r = Nv$.

How to *best* solve (*) is open to question. Mathematically, we can solve

$$L'Lu = S'b, \qquad W'Wv = N'b$$

for u and v, but computation of the "Gram matrices" $L'L$ and $W'W$ can certainly be avoided. A direct approach would be to simply factor $[L \;\; W]$ using Gauss factorization. In many applications $n-\rho$ is small (0, 1 or 2, say). Note that only one of p or r needs to be computed since, e.g., $r = b - p$.

2. p and r are now fixed. Let $y \in \mathcal{R}(A)$. Since $p \in \mathcal{R}(A)$, $r \in \mathcal{N}(A')$ and $\mathcal{N}(A') \perp \mathcal{R}(A)$ then $r'(p-y) = 0$. Thus, by the Pythagorean theorem,

$$\begin{aligned} \|b-y\|^2 &= \|b-p+p-y\|^2 \\ &= \|r+(p-y)\|^2 \\ &= \|r\|^2 + \|p-y\|^2 \\ &\geq \|r\|^2, \end{aligned}$$

with equality if and only $y = p$.

3. Since the columns of R form a basis for $\mathcal{R}(A')$, every $x^* \in \mathcal{R}(A')$ can be written uniquely as

$$x^* = Rs, \qquad s \in \mathbb{R}^\rho\,.$$

Since the columns of M form a basis for $\mathcal{N}(A)$, every $z \in \mathcal{N}(A)$ can be written uniquely as

$$z = Mt, \qquad t \in \mathbb{R}^{m-\rho}\,.$$

Thus the equation $x^* + z = x_p$ is equivalent with the linear system

$$Rs + Mt = x_p\,.$$

Moreover, since $R = PK$ ($K = U'$) and $M = PV$ with P a permutation matrix, this system is equivalent with

$$Ks + Vt = P'x_p ,$$

that is with

$$\begin{bmatrix} K_1 & V_2 \\ K_2 & I \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} = P'x_p . \qquad (**)$$

As in #1 the matrix

$$\begin{bmatrix} L & V \end{bmatrix} = \begin{bmatrix} K_1 & V_2 \\ K_2 & I \end{bmatrix}$$

is nonsingular and (**) could be solved by solving the positive definite systems

$$K'Ks = R'x_p , \qquad V'Vt = M'x_p .$$

Thus s and t are uniquely determined (by x_p) and then so are $x^* = Rs$ and $z = Mt$. Again, only one of x^* or z needs to be computed since, e.g., $z = x_p - x^*$.

4. Let x_p solve $Ax = p$. Since $Az = 0$ then $x^* = x_p - z$ is *also* a particular solution of $Ax = p$. Let x be an *arbitrary* solution of $Ax = p$. Then

$$x = x^* + M\hat{s}, \qquad \hat{s} \in F^{\rho} .$$

Since $Ms \in \mathcal{N}(A)$ and $x^* \in \mathcal{R}(A') \perp \mathcal{N}(A)$ then, by the Pythagorean theorem,

$$\begin{aligned} \| x \|^2 &= \| x^* + M\hat{s} \|^2 \\ &= \| x^* \|^2 + \| M\hat{s} \|^2 \\ &\geq \| x^* \|^2 , \end{aligned}$$

with equality if and only if $\hat{s} = 0$, that is $x = x^*$.

■

Definition. In the "FTLA":

- p is *the orthogonal projection* of b onto $\mathcal{R}(A)$.
- r is *the* orthogonal projection of b onto $\mathcal{N}(A')$,
- x^* is *the* orthogonal projection of *every* solution of $Ax = p$ onto $\mathcal{R}(A')$.

One algorithm gfss (*first draft*, $P = I_m$, $Q = I_n$).

1. solve $L'Lu = L'b$ for u,

2. compute $p = Lu$, $r = b - p$,

3. solve $U_1 x_1 = u$ for x_1 and put $x_p = \begin{array}{cc} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} & \begin{array}{c} \rho \\ m-\rho \end{array} \end{array}$.

4. solve $UU's = Ux_p$ for s,

5. compute $x^* = U's$.

Note that $Ax = p \Leftrightarrow LUx = Lu \Leftrightarrow Ux = u$, justifying step 3. However, note also that $Ux_p = U_1 x_1 = u$, so step 3 and the multiplication Ux_p in step 4 can be avoided!

One algorithm gfss (second draft, $P = I_m$, $Q = I_n$).

1. solve $L'Lu = L'b$ for u,

2. compute $p = Lu$, $r = b - p$,

3. solve $UU's = u$ for s,

4. compute $x^* = U's$.

Observe that the *residual vector* $r = b - Ax^*$ is computed *before* x^* is computed!

As we indicated above, there should be better algorithms, but our code gfss is not bad for starters. It uses Cholesky factorization of $L'L$ and UU' (gfcc).

Formally, if

$$Q'AP = LU = LK'$$

is an LU factorization then, for every $b \in \mathbb{F}^n$, the shortest solution of the least squares problem $\| b - Ax \|^2 = \text{minimum}$ is

$$x^* = A^I b$$

with

$$\begin{aligned} A^I &= PU'(UU')^{-1}(L'L)^{-1}L'Q' \\ &= PK(K'K)^{-1}(L'L)^{-1}L'Q' \end{aligned}$$

one representation of the *Moore-Penrose pseudoinverse* of A. A much simpler representation of A^I can be given as a consequence of the "svd" ("singular value decomposition") of A. However, computation of the svd involves an eigenvalue problem. Here we have used only Gauss factorization.

In any case A^I *should not be computed as a means for solving the (generalized) least squares problem.* It is (merely) a theoretical entity!

Example.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -3 & -2 \\ 1 & 1 & 2 \\ 1 & -2 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

On page 15 we found that $Ax = b$ is not solvable, exactly. On pages 5-8 we factored

$$A = LU = \begin{bmatrix} 1 & \\ 1 & 1 \\ 1 & -1/3 \\ 1 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ & -3 & -3 \end{bmatrix}$$

with no pivoting. First we have

$$L'L = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & \\ 1 & 1 \\ 1 & -1/3 \\ 1 & 2/3 \end{bmatrix} = \begin{bmatrix} 4 & 4/3 \\ 4/3 & 14/9 \end{bmatrix}.$$

$$UU' = \begin{bmatrix} 1 & 0 & 1 \\ & -3 & -3 \end{bmatrix} \begin{bmatrix} 1 & \\ 0 & -3 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ -3 & 18 \end{bmatrix}.$$

We factor, by Gauss,

$$L'L =: \begin{bmatrix} 1 & \\ 1/3 & 1 \end{bmatrix} \begin{bmatrix} 4 & 4/3 \\ & 10/9 \end{bmatrix},$$

$$UU' = \begin{bmatrix} 1 & \\ -3/3 & 1 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ & 27/2 \end{bmatrix},$$

and compute

$$L'b = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 11/3 \end{bmatrix}.$$

1. Solve $L'Lu = L'b$ for u:

$$\begin{bmatrix} 1 & \\ 1/3 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 4 & 4/3 \\ & 10/9 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}}_{=:\ c = \begin{bmatrix} 10 \\ 1/3 \end{bmatrix}} = \begin{bmatrix} 10 \\ 11/3 \end{bmatrix}$$

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 24 \\ 3 \end{bmatrix} \frac{1}{10}.$$

2. Compute

$$p = Lu = \begin{bmatrix} 1 & \\ 1 & 1 \\ 1 & -1/3 \\ 1 & 2/3 \end{bmatrix} \begin{bmatrix} 24 \\ 3 \end{bmatrix} \frac{1}{10} = \begin{bmatrix} 24 \\ 27 \\ 23 \\ 26 \end{bmatrix} \frac{1}{10},$$

$$r = b - p = \begin{bmatrix} -14 \\ -7 \\ 7 \\ 14 \end{bmatrix} \frac{1}{10} = \begin{bmatrix} -2 \\ -1 \\ 1 \\ 2 \end{bmatrix} \frac{7}{10},$$

and

$$\| r \|^2 = \frac{49}{10}, \qquad \| r \| = \frac{7}{\sqrt{10}}.$$

3. Solve $UU's = u$ for s:

$$\begin{bmatrix} 1 & \\ -3/2 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 2 & -3 \\ & 27/2 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix}} = \begin{bmatrix} 24 \\ 3 \end{bmatrix} \frac{1}{10}$$

$$=: \ c = \begin{bmatrix} 24 \\ 39 \end{bmatrix} \frac{1}{10}.$$

$$s = \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix} = \begin{bmatrix} 49/2 \\ 13/3 \end{bmatrix} \frac{1}{15}.$$

4. Compute

$$x^* = U's = \begin{bmatrix} 1 & \\ 0 & -3 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 49/2 \\ 13/3 \end{bmatrix} \frac{1}{15}$$

$$= \begin{bmatrix} 49/2 \\ -13 \\ 23/2 \end{bmatrix} \frac{1}{15} = \begin{bmatrix} 49 \\ -26 \\ 23 \end{bmatrix} \frac{1}{30}.$$

Check. On page 14 we computed the "basis matrix"

$$M = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

for $\mathcal{N}(A)$. We have

$$M'x^* = \begin{bmatrix} -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 99 \\ -26 \\ 23 \end{bmatrix} \frac{1}{30} = 0.$$

Problem 13.

Find the shortest solution of the least squares problem for

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix},$$

as above, including the check.

Problem 14.

Let $Q'AP = LU$ with $p = \begin{bmatrix} p(1) & p(2) & \ldots & p(m) \end{bmatrix}$ the permutation associated with P. Show that the columns $\{a_{p(1)}, a_{p(2)}, \ldots, a_{p(p)}\}$ form a basis for $\mathcal{R}(A)$.

```
function [L,U,g] = gfpn(A,dummy)

%   [L U g] = gfpn(A,dummy):
%
%   GAUSS FACTORIZATION with NO PIVOTING for size.  Tries to factor A = LU
%   with L unit lower trapezoidal and U upper trapezoidal.
%
%   g is the growth factor.  It is the ratio of the maximum absolute value
%   of the elements of all Gauss transforms, including A itself, to the
%   maximum absolute value of the elements of A.
%
%   Displays successive tableaus if nargin, the number of input arguments,
%   is two and cols A < 8.  This code provides a nice introduction to matlab
%   programming.
%
%   gfpn is not suitable for general use since it is not numerically stable:
%   For certain important classes of matrices, however, it can be shown that
%   pivoting for size is not strictly necessary.  For such problems gfpn is
%   quite useful.  See gfmarkov for instance.

%   Copyright (c) 6 April 1991 by Bill Gragg.  All rights reserved.
%   Revised 3 October 1994.

%   gfpn calls no extrinsic functions.


%   Here is a streamlined "production type" code, without comments,
%   displays, or computation of the growth factor.  It is easier to see the
%   essence of the code here, without the frills.  The lines '% begin gfpn'
%   and '%end gfpn' are part of my own style, a carryover from algol, the
%   first mathematically precise programming language.  Three space indent-
%   ation is fairly standard.  The vertical spacing is rather subjective.

%   function [L,U] = gfpn(A)
%   begin gfpn
%      [n m] = size(A);   k = 0;
%      for k = 1:min(m,n)
%         if k < n
%            j = k+1:n;  a = A(j,k);
%            if A(k,k) ~= 0
%               a = a/A(k,k);   A(j,k) = a;
%            else
%               if any(a ~= 0)
%                  error('Attempt to divide by zero in pivot position.')
%               end
%            end
%            if k < m
%               i = k+1:m;
%               A(j,i) = A(j,i) - a*A(k,i);
%            end
%         end
%      end
%      if k > 0
%         L = A(:,1:k);   U = A(1:k,:);   L(1,1) = 1;   z = zeros(k,1);
%         for i = 1:k-1
%            o = z(1:i);   L(1:i+1,i+1) = [o; 1];
%            j = k - 1;    U(j+1:k,j) = o;
%         end
%      end
```

```
%       end gfpn

%       Here is the actual code.  It is more complicated because of all its
%       options.

%       begin gfpn

%          Find the numbers of columns (m) and rows (n) of A and initialize the
%          stage counter k.

           [n m] = size(A);   k = 0;

%          Set logical variables.

           display = nargin > 1  &  m < 8;
           growth  = display  |  ( nargout > 2 );

%          Compute the scale factor a and initialize the growth factor g, if
%          required.

           if growth
              s = max(abs(A(:)));   g = s;
           end

%          Display k, A and s, if required.

           if display

              title = 'Pauses after each stage.  Press any key to continue.';

              format compact,   format short,   disp(' '),   disp(title)
              disp(' '),        k,              A,           s
              pause

           end

%          "Do Gauss".  The array A should now be thought of as a "tableau", or
%          workspace, not as a matrix.

           for k = 1:min(m,n)

%             Compute the new tableau.  Here matlab's matrix operations allow
%             one to avoid the use of a nest of two for loops.  Loops are
%             extremely slow in matlab.

              if k < n

                 j = k+1:n;   a = A(j,k);

                 if A(k,k) ~= 0

%                   Divide elements below the pivot by the pivot.

                    a = a/A(k,k);   A(j,k) = a;

                 else

                    if any(a ~= 0)
```

```
            error('Attempt to divide by zero in pivot position.')
         end

      end

      if k < m

%        Update the "Gauss transform" part of the tableau.

         i = k+1:m;
         A(j,i) = A(j,i) - a*A(k,i);

%        Update the growth factor, if required.

         if growth
            t = max(max(abs(A(j,i))));   g = max(g,t);
         end

      end

   end

%  Display the stage counter k, the modified tableau and the current
%  growth, if required.

   if display

      k,   A

      if s > 0
         h = g;   g = g/s,   g = h;
      else
         g = 1
      end

      pause

   end

end

%  Prepare the output.  L is the lower trapezoidal part of the final
%  tableau with its diagonal elements replaced by ones.  U is the upper
%  trapezoidal part of the final tableau.

if k > 0

   L = A(:,1:k);   U = A(1:k,:);   L(1,1) = 1;   z = zeros(k,1);

   for i = 1:k-1
      o = z(1:i);   L(1:i+1,i+1) = [o; 1];
      j = k - i;    U(j+1:k,j) = o;
   end

end

if growth  &  s > 0
   g = g/s;
end
```

```
%       end gfpn

%       Approximate total flops [m = cols(A), n = rows(A)]:
%          Real case:       TBC
%          Complex case:    TBC
```

```
        function [L,U,p,q,g] = gfpc(A,dummy)

%       [L U p q g] = gfpc(A,dummy):
%
%       GAUSS FACTORIZATION with COMPLETE PIVOTING for size.  Factors Q'AP = LU
%       with P and Q permutation matrices, L unit lower trapezoidal with r
%       columns and U upper trapezoidal with nonzero diagonal elements and r
%       rows.  p and q are the permutations (row vectors) associated with P and
%       Q, g is the growth factor, and r is a classical determination of a
%       "numerical rank" of A.
%
%       Displays intermediate output with pauses if nargin = 2 and cols(A) < 8.
%       Press any key to continue the computation.

%       Copyright (c) 6 April 1991 by Bill Gragg.  All rights reserved.
%       Revised 10 April 1994.

%       gfpc calls no extrinsic functions.


%       begin gfpc

%          Initialize p, q and g.

           [n m] = size(A);    p = 1:m;    q = 1:n;    g = 0;

%          Set up displays and display initial tableau, if required.

           display = nargin > 1  &  m < 8;

           if display

              format compact,   format short

              title1 = 'Initial tableau / p / q';
              title2 = 'Pivot indices (t,s) / Pivot A(t,s) / Current growth g';
              title3 = 'Tableau after interchanges / p / q';
              title4 = 'Tableau after computation / p / q';

              disp(' '),     disp(title1),    disp(A)
              disp(p),       disp(q),         pause

           end

%          Perform the Gauss factorization.

           for r = 1:min(m,n)

%             Find the pivot and update the "local growth" g.

              [b j] = max(abs(A(r:n,r:m)));    [c s] = max(b);    t = j(s);
              s = r + s - 1;                   t = r + t - 1;     g = max(g,c);

%             Adjust for matlab's "nonuniform" max function.

              if r == n
                 s = t;  t = r;
              end
```

```
%         Set scale factor and tolerance if r = 1.

          if r < 2
             a = g;    tol = min(m,n)*a;
          end

%         Terminate loop, if possible.

          if tol + c == tol
             r = r - 1;    break
          end

%         Display pivot and local growth information, if required

          if display
             disp(sprintf('Stage %g',r)),    disp(title2),    disp([t s])
             disp(A(t,s)),                   disp(g/a),       pause
          end

%         Perform the interchanges.

          p([r s]) = p([s r]);  A(:,[r s]) = A(:,[s r]);
          q([r t]) = q([t r]);  A([r t],:) = A([t r],:);

%         Display tableau after interchanges, if required.

          if display
             disp(title3),    disp(A),    disp(p),    disp(q),    pause
          end

%         This completes the half-step.  Now do the computation.

          if r < n

             j = r+1:n;    A(j,r) = A(j,r)/A(r,r);

             if r < m
                i = r+1:m;
                A(j,i) = A(j,i) - A(j,r)*A(r,i);
             end

          end

%         Display new tableau, if required.

          if display  &  ( r < n  |  m < n )
             disp(title4),    disp(A),    disp(p),    disp(q),    pause
          end

       end

%      Prepare the output.  L is the lower trapezoidal part of A with its
%      diagonal elements replaced by ones.  U is the upper trapezoidal part
%      of A.  Scale g by a if r > 0;  otherwise g := 1.

       if r > 0

          L = A(:,1:r);    U = A(1:r,:);    L(1,1) = 1;    z = zeros(r,1);

          for i = 1:r-1
```

```
            o = z(1:i);    L(1:i+1,i+1) = [o; 1];
            j = r - i;     U(j+1:r,j) = o;
         end

         g = g/a;

      else

         g = 1

      end

%     end gfpc


%     Approximate total flops [m = cols A, n = rows A]:
%        Real case:     TBC
%        Complex case:  TBC
```

```
        function [L,U,p,q,g] = gf(A)

%       [L U p q g] = gf(A):
%
%       GAUSS FACTORIZATION with COMPLETE PIVOTING for size.  Factors Q'AP = LU
%       with P and Q permutation matrices, L unit lower trapezoidal with r
%       columns, and U upper trapezoidal with nonzero diagonal elements and r
%       rows.  p and q are the permutations (row vectors) associated with P and
%       Q, g is the growth factor, and r is classical determination of a
%       "numerical rank" of A.
%
%       This is our matlab "production code".  It is the same as gfpc, but with
%       no comments or displays.

%       Copyright (c) 8 June 1991 by Bill Gragg.  All rights reserved.
%       Revised 10 April 1994.

%       gf calls no extrinsic functions.


%       begin gf

          [n m] = size(A);    p = 1:m;    q = 1:n;    g = 0;

          for r = 1:min(m,n)

             [b j] = max(abs(A(r:n,r:m)));    [c s] = max(b);    t = j(s);
             s = r + s - 1;                   t = r + t - 1;     g = max(g,c);

             if r == n
                s = t;   t = r;
             end

             if r < 2
                a = g;   tol = min(m,n)*a;
             end

             if tol + c == tol
                r = r - 1;   break
             end

             p([r s]) = p([s r]);  A(:,[r s]) = A(:,[s r]);
             q([r t]) = q([t r]);  A([r t],:) = A([t r],:);

             if r < n

                j = r+1:n;   A(j,r) = A(j,r)/A(r,r);

                if r < m
                   i = r+1:m;   A(j,i) = A(j,i) - A(j,r)*A(r,i);
                end

             end

          end

          if r > 0

             i = 1:r;       L = A(:,i);       U = A(i,:);
```

```
      L(1,1) = 1;    z = zeros(r,1);

      for i = 1:r-1
         o = z(1:i);    L(1:i+1,i+1) = [o; 1];
         j = r - i;     U(j+1:r,j) = o;
      end

      g = g/a;

   else

      g = 1;

   end

%      end gf
```

PROBLEM SET: FORWARD AND BACKWARD SOLUTION OF TRIANGULAR SYSTEMS. MATRIX MULTIPLICATION.

In problems 1–4, solve the given triangular system, $Lc = b$ or $Ux = c$, by hand.

Problem 1.

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 2 \\ 1 \\ 0 \\ -2 \end{bmatrix}.$$

Answer: $c = [2,\ 3,\ 5,\ 8]'$.

Problem 2.

$$U = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{bmatrix}, \qquad c = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 8 \end{bmatrix}.$$

Answer: $x = [1,\ 1,\ 1,\ 1]'$.

Problem 3.

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 1 & 7 & 6 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 \\ -2 \\ -10 \\ -44 \end{bmatrix}.$$

Answer: $c = [0,\ -2,\ -4,\ -6]'$.

Problem 4.

$$U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 6 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 \\ -2 \\ -4 \\ -6 \end{bmatrix}.$$

Answer: $x = [1, \ -1, \ 1, \ -1]'$

Problem 5.

Combine the results of problems 1 and 2. Compute $A := LU$ and check that $Ax = b$, by hand.
Answer: $A = \text{mxwilkinson}(4)$. See Gauss factorization problem set.

Problem 6.

Combine the results of problems 3 and 4. Compute $A := LU$ and check that $Ax = b$, by hand.
Answer: $A = \text{mxvandermonde}([1 \ \ 2 \ \ 3 \ \ 4])$. See Gauss factorization problem set.

PROBLEM SET: GAUSS FACTORIZATION

Factor the following structured matrices A into $A = LU$, by hand, using the tableau, and check your work by matrix multiplication. You will better appreciate computers after doing these problems, and they really are interesting! You can find out the complete answers by executing [L U g] = gfpn(A, 0) in matlab. The required special matrix codes are in stewart/home/ma1043/mfiles.

"Do Gauss" – NO PIVOTING. Note well the growth factors g.

Problem 1.

A 4 by 4 matrix due to Wilkinson:

$$W = \text{mxwilkinson}(4),$$

$$W = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}.$$

What is the growth factor for mxwilkinson(n)?

Problem 2.

A Hadamard matrix of order 4:

$$H = \text{mxhadamard}(4),$$

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Problem 3.

"Pascal's matrix" of order 5:

$$P = \text{mxpascal}(5),$$

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 5 & 15 & 35 & 70 \end{bmatrix}.$$

This is an example of a *Cholesky factorization.* We have $U = L$ so that $A = LL'$, with L having *positive diagonal elements.* Also, the *Cholesky factor L* is a part of Pascal's triangle, written another way.

Problem 4.

The (tridiagonal) "negative second difference matrix" of order 5:

$$T = \text{mxtsd}(5),$$

$$T = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

(The second most important matrix in the Whole Wide World)

Note that the factorization process is very "cheap" for tridiagonal matrices. How cheap!

You can modify the factorization $T = LU$ in a simple way, by an "interior diagonal scaling," to get another Cholesky factorization $T = R'R$, with R upper triangular.

The diagonal elements of U, the pivots, are all positive. Let D be the diagonal matrix formed from their (positive) square roots. Then $T = LDD^{-1}U = R'R$ with $R' := LD$ and $R = D^{-1}U$. Thus one *multiplies* the *columns* of L by the square roots of the pivots and, to adjust for this, *divides* the *rows* of U by these square roots.

Problem 5.

The "negative periodic second difference matrix" of order 5:

$$T = \text{mxpsd}(5),$$

$$T = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

The last diagonal element of U will be zero! This can also be modified slightly to be a Cholesky factorization, $T = R'R$, with the last diagonal element of R zero.

Problem 6.

The "min matrix" of order 5:

$$M = \text{mxmin}(5),$$

$$M = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}.$$

Another Cholesky factorization!

Problem 7.

The "max matrix" of order 5:

$$M = \text{mxmax}(5),$$

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 3 & 4 & 5 \\ 3 & 3 & 3 & 4 & 5 \\ 4 & 4 & 4 & 4 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{bmatrix}.$$

Problem 8.

A 4 by 4 Vandermonde matrix built from the "abscissas" 1, 2, 3, 4.

$$V = \mathrm{mxvandermonde}([1\ \ 2\ \ 3\ \ 4]),$$

$$V = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{bmatrix}. \quad \text{(V is of } \textit{theoretical} \text{ interest)}$$

Problem 9.

The 3 by 3 Hilbert matrix:

$$H = \mathrm{rats}(\mathrm{mxhilbert}(3)),$$

$$H = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}.$$

Problem 10.

The "idft matrix" of order 4 (the *most important* matrix in the Whole Wide World):

$$W = \mathrm{mxidft}(4),$$

$$W = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix} \quad (i := \mathrm{sqrt}(-1))$$

Complex matrices rarely arise in practice but, when they do, they seem to be important. mxidft(n) is probably the most important matrix of all time, the matrix used in the *fast Fourier transform*. Typical orders are $n = 1024$ and $n = 4096$!. Here we use the case $n = 4$ as a, fairly massive (!), drill in complex arithmetic.

Complete pivoting for size or, more precisely, pivoting to *prevent growth.*

This kind of pivoting should *not* be confused with the term "pivoting" that is used in the field of linear constrained optimization (linear programming), nor should the term "programming" which is used in that field be confused with the programming of computers! Confusing, isn't it?

Problem 11.

Factor the Wilkinson matrix of problem 1 using complete pivoting: $Q'AP = LU$. What is the growth factor now? Answer: $g = 2$!

Problem 12.

Factor $Q'AP = LU$ for the following matrix A in *three* ways, using no pivoting, complete pivoting, and any *other* pivot scheme you choose.

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 5 \end{bmatrix}.$$

Pivots are, by definition, not zero. How many pivots are there in *each* case? We will ultimately show that, for a given matrix A, *every* pivot scheme will always find the *same* number of pivots, and this no matter how large the matrix A!

These are examples of the

LU Theorem. Let A be n by m with $A \neq O$. There are permutation matrices P and Q, and an integer r with $1 <= r <= \min(m, n)$, so that

$$Q'AP = LU\ ,$$

with L unit lower trapezoidal with r columns, and U upper trapezoidal with nonzero diagonal elements (and r rows).

Now you have "paid your dues" so you can use matlab.

Problem 13.

Factor the 8 by 8 versions of the matrices in problems 1–11 with matlab. Use all three codes: gfpn, gfppr and gfpc. *Do not output the factorizations!* Compare the growth factors g obtained by these three pivot strategies. Check the factorizations by displaying the respective scaled errors

$$\text{en} = \text{norm}\,(A - L*U)/a$$
$$\text{ep} = \text{norm}\,(A(q, :) - L*U)/a$$
$$\text{ec} = \text{norm}\,(A(q, p) - L*U)/a$$

where

$$a = \text{norm}\,(A)\,.$$

Simultaneously, record the *condition numbers*, cond A, of these matrices. For instance the "one liner" cond(mxhilbert(8)) gives the condition number of the 8 by 8 Hilbert matrix.

Repeat with $n = 8$ replaced by $n = 16$, or do these simultaneously.

(Not so) roughly speaking one can expect to lose log10(cond$A$) decimal digits of accuracy when solving $Ax = b$ for x on a computer, for square matrices A. We always have cond A >= 1. A is *ill-conditioned* if cond A is *large.* One might think that large growth and ill-conditioning are related. The following example shows that this is not true.

We have $g = 2^{n-1}$ for W = mxwilkinson(n) and partial pivoting, but $g = 2$ if complete pivoting is used, it appears. (One can prove this!) What is cond(mxwilkinson(200))? How long does it take matlab to compute it? How about $n = 500$?

Problem 14.

Factor the matrix in problem 12 by using matlab and gfpc, and check that ec is of the order of magnitude of the machine precision eps.

Problem 15.

Execute "eps" and "binrep(eps)", or just "br(eps)". Then execute "eps = machprec(0)", to replace eps by its "correct" value. We won't have to know the fine details concerning eps.