Start with any floating point number $x^* \in F_{64}$ → round → to the nearest decimal with m significant digits → round → to the nearest floating point number $y^* \in F_{64}$

Question what is the smallest m such that $x^* = y^*$ for all $x^* \in F_{64}$ with the above path of roundings.

17

Start with any decimal number x with m significant digits → round → to the nearest floating point number → round → to the nearest decimal with m significant digits y.

Question what is the largest m such that $x = y$ for all x with m significant digits.

15

So if $m = 16$ there is a floating number $x^* \in F_{64}$ such that

$x^* \longrightarrow$ 16 sig. dec digits $\longrightarrow y^* \in F_{64}$ that $x^* \neq y^*$

also there is x with 16 significant digits such that

$x \longrightarrow$ 64-bit floating point $\longrightarrow y$ with 16 sig dig. that $x \neq y$.

Extra credit. find an example for each round trip failure when $m = 16$

Cholesky Decomposition:

Fact: If U_1 is upper triangular, U_2 is upper triangular, then U_1U_2 is upper triangular.

Fact: If U_1 is lower triangular, U_2 is lower triangular, then U_1U_2 is lower triangular.

```
julia> include("myLU.jl")
myLU (generic function with 1 method)

julia> Z=rand(3,3)
3×3 Matrix{Float64}:
 0.612015  0.659229  0.525741
 0.573422  0.214699  0.348855
 0.201189  0.55424   0.244285

julia> A=Z*Z'
3×3 Matrix{Float64}:
 1.08555   0.675886  0.616931
 0.675886  0.496608  0.319581
 0.616931  0.319581  0.407333
```

Positive definite means $Ax \cdot x \geq 0$ for every $x \in \mathbb{R}^n$

$$Ax \cdot x = ZZ^T x \cdot x = (ZZ^T x)^T x = x^T Z^{TT} Z^T x = x^T Z Z^T x$$

($Z^T x)^T$

$$= (Z^T x)^T Z^T x = Z^T x \cdot Z^T x = \|Z^T x\|_2^2 \geq 0$$

since norms are never negative..

```
julia> L1,U1=myLU(A)
([1.0 0.0 0.0; 0.6226220038668691 1.0 0.0; 0.5683134987379813 -0
.8515223078177523 1.0], [1.0855476354454712 0.6758858440740007 0
.6169313747467584; -1.1102230246251565e-16 0.07578707536225027 -
0.06453438531522132; -9.453796721212185e-17 0.0 0.00177051810874
92236])

julia> L1
3×3 Matrix{Float64}:
 1.0        0.0       0.0
 0.622622   1.0       0.0
 0.568313  -0.851522  1.0

julia> U1
3×3 Matrix{Float64}:
  1.08555      0.675886    0.616931
 -1.11022e-16  0.0757871  -0.0645344
 -9.4538e-17   0.0         0.00177052
```

these are the factors α in the elimination steps $r_i \leftarrow r_i - \alpha r_j$

```
julia> A
3×3 Matrix{Float64}:
 1.08555   0.675886  0.616931
 0.675886  0.496608  0.319581
 0.616931  0.319581  0.407333

julia> L1*U1
3×3 Matrix{Float64}:
 1.08555   0.675886  0.616931
 0.675886  0.496608  0.319581
 0.616931  0.319581  0.407333

julia> using LinearAlgebra

julia> norm(A-L1*U1)
0.0
```

- Using the LU code in the book translated to Julia...

```
julia> D1=Diagonal(U1)
3×3 Diagonal{Float64, Vector{Float64}}:
 1.08555   ⋅          ⋅
  ⋅       0.0757871   ⋅
  ⋅        ⋅         0.00177052

julia> U1p=D1^-1*U1
3×3 Matrix{Float64}:
  1.0           0.622622   0.568313
 -1.46492e-15   1.0       -0.851522
 -5.33957e-14   0.0        1.0
```

⋅ means the number is not stored in memory and assumed to be zero

```
julia> U1p=UpperTriangular(D1^-1*U1)
3×3 UpperTriangular{Float64, Matrix{Float64}}:
 1.0  0.622622   0.568313
  ⋅   1.0       -0.851522
  ⋅    ⋅         1.0
```

zeros

```
julia> L1*D1*U1p
3×3 Matrix{Float64}:
 1.08555   0.675886  0.616931
 0.675886  0.496608  0.319581
 0.616931  0.319581  0.407333

julia> A
3×3 Matrix{Float64}:
 1.08555   0.675886  0.616931
 0.675886  0.496608  0.319581
 0.616931  0.319581  0.407333
```

```
julia> U1p
3×3 UpperTriangular{Float64,
 1.0  0.622622   0.568313
  ⋅   1.0       -0.851522
  ⋅    ⋅         1.0

julia> L1
3×3 Matrix{Float64}:
 1.0        0.0       0.0
 0.622622   1.0       0.0
 0.568313  -0.851522  1.0
```

Note $L_1 = U_1^T$

Assume A is invertible.

lower triangular ones on diagonal ← diagonal → upper triangular ones on the diagonal.

Note that if $A = L_1 D_1 U_1$ and $A = L_2 D_2 U_2$ then $L_1 = L_2$ and $D_1 = D_2$ and $U_1 = U_2$. That is, the LDU factorization is unique.

Fact: If L is lower triangular with ones on the diagonal then L^{-1} is lower triangular with ones on the diagonal

Why? $L = I - N$ $\quad N = I - L$

ones on diagonal

```
julia> N=I-L1
3×3 Matrix{Float64}:
  0.0        -0.0       -0.0
 -0.622622    0.0       -0.0
 -0.568313    0.851522   0.0
```

strictly lower triangular

```
julia> N*N
3×3 Matrix{Float64}:
  0.0        -0.0  0.0
  0.0         0.0  0.0
 -0.530177    0.0  0.0
```

```
julia> N^3
3×3 Matrix{Float64}:
 0.0  -0.0  0.0
 0.0   0.0  0.0
 0.0   0.0  0.0
```

So in general if $L \in \mathbb{R}^{n \times n}$ and $L = I - N$ then $N^n = 0$

$$S = I + N + N^2 + \cdots + N^{n-1} + N^n + N^{n+1}$$

all are zero...

subtract $\quad NS = N + N^2 + N^3 + \cdots + N^n + N^{n+1} +$

all are zero...

$$LS = (I - N)S = I$$

lower triangular.

which means $S = I + N + \cdots + N^{n-1}$ is the inverse of $I - N = L$

the inverse has ones on the diagonal.

Assume A is invertible ... then D^{-1} has to exist

$A = L_1 D_1 U_1$ and $A = L_2 D_2 U_2$

$L_1 D_1 U_1 = L_2 D_2 U_2$

$L_2^{-1} L_1 D_1 = D_2 U_2 U_1^{-1}$

lower triangular — upper triangular

$\Rightarrow$ that $D_1 = D_2$

$L_2^{-1} L_1 = D U_2 U_1^{-1} D^{-1}$

lower triangular with ones on diagonal — upper triangular

So $L_2^{-1} L_1 = I$ and $L_2 = L_1$.