Solving Linear Equations

Dept of Mechanical & Industrial Engineering
University of Iowa

Contents

- Preliminaries
  - o Elementary Matrix Operations
  - o Elementary Matrices
  - o Echelon Matrices
  - o Rank of Matrices
- Elimination Methods
  - o Gauss Elimination
  - o Gauss-Jordan Elimination
- Factorization Methods
  - o Product Form of Inverse (PFI)
  - o LU Factorization
  - o LDL^T Factorization
  - o Cholesky LL^T Factorization

Elementary Row Operations

- Multiply any row of the matrix by a (positive or negative) scalar
- Add to any row a scalar multiple of another row
- Interchange two rows of the matrix

(Strictly speaking, the third is not "elementary", because it can be accomplished by a sequence of the other two row operations!)

Elementary Column Operations

- Multiply any column by a (positive or negative) scalar
- Add to any column a scalar multiple of another column
- Interchange two columns of the matrix

Equivalence of Matrices

Matrix **A** is ***equivalent*** to matrix B ($A \sim B$) if B is the result of a sequence of elementary row &/or column operations on **A**.

If only row operations are used, then **A** is ***row-equivalent*** to B

If only column operations are used, then **A** is ***column-equivalent*** to B

Echelon Matrix

--an mxn matrix with the properties

- each of the first k ($0 \le k \le m$) rows has some nonzero entries, and the remaining m-k rows consist only of zeroes
- the first nonzero entry in each of the first k rows is a "1"
- in each of the first k rows, the number of zeroes preceding the leading "1" is smaller than it is in the next row

ECHELON MATRIX

Example

$$\begin{bmatrix} 1 & 5 & 0 & 3 & -1 & 2 & 8 \\ 0 & 0 & 1 & -1 & 2 & 5 & 0 \\ 0 & 0 & 0 & 1 & 3 & 1 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} k=3 \quad = \textit{rank}$$

Note: every matrix is row-equivalent to some echelon matrix.

Theorem

If A is equivalent to B, then the rank of A equals the rank of B.

RANK: size of the largest (square) nonsingular submatrix

Elementary Matrices

An ***elementary matrix*** E is the result of performing an elementary operation on an identity matrix.

Example

(Elementary row operation: add −2 times first row to third row)

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

Multiplication by an Elementary Matrix

pre-multiplication by elementary matrix

If E is an mxm elementary matrix and A is an mxn matrix, then EA equals the result of performing the same elementary ***row*** operation on matrix A.

Example:

add −2 times first row to third row

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 4 \\ 5 & 1 & 3 & -1 \\ 4 & 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 4 \\ 5 & 1 & 3 & -1 \\ 0 & 5 & 1 & -6 \end{bmatrix}$$

If E is an mxm elementary matrix and A is an mxn matrix, then AE equals the result of performing the same elementary ***column*** operation on matrix A.

Example:

add −2 times third ***column*** *to first* ***column***

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

result of subtracting twice third column from first

post-multiplication by elementary matrix

Calculation of Matrix Inverse

To compute A^{-1}, augment the matrix A on the right by the appropriate identity matrix $[A \mid I]$, and perform elementary row operations on this matrix to obtain $[I \mid P]$. Then $P = A^{-1}$

Calculation of Matrix Inverse

Example:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & -5 & 3 \\ 0 & 1 & 0 & 3 & 3 & -2 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right]$$

and so

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

Pivot

Pivot operation on row **r**, column **s**
i.e., element A_r^s of mxn matrix A:

A sequence of elementary row operations:

- For i=1,2,...m but i≠r:
  add $-A_i^s / A_r^s$ times row r to row i
- Multiply row r by the scalar $1/A_r^s$

Effect: column s will consist of zeroes, with the exception of a "1" in row r.

Warning: this is not the only sequence of elementary row operations having this effect!

Pivot

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

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

A pivot!

$R_1 \leftarrow R_1 - 1/3\, R_3$
$R_2 \leftarrow R_2 - 1/3\, R_3$
$R_3 \leftarrow 1/3\, R_3$

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

***Not** a pivot!*

$R_1 \leftarrow R_1 - R_2$
$R_2 \leftarrow R_2 - 1/3\, R_3$
$R_3 \leftarrow 1/3\, R_3$

Pivot Matrix

A pivot matrix corresponding to a pivot on row r, column s of a matrix A is the result of performing the same elementary row operations on the mxm identity matrix.

A pivot matrix is the product of elementary matrices!

Pivot Matrix

Differs from the mxm identity matrix only in column r

$$\begin{bmatrix} 1 & 0 & \cdots & -\frac{A_1^s}{A_r^s} & \cdots & 0 & 0 \\ 0 & 1 & \cdots & -\frac{A_2^s}{A_r^s} & \cdots & 0 & 0 \\ & & \ddots & & & & \\ 0 & 0 & \cdots & \frac{1}{A_r^s} & \cdots & 0 & 0 \\ 0 & 0 & \cdots & -\frac{A_{m-1}^s}{A_r^s} & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -\frac{A_m^s}{A_r^s} & \cdots & 0 & 1 \end{bmatrix}$$

Pivot Matrix

To store a pivot matrix, we need not store the entire matrix, but only

- the number (r) of the pivot row
- column #r of the pivot matrix (the ***eta*** vector)

$$\eta = \left[-\frac{A_1^s}{A_r^s}, -\frac{A_2^s}{A_r^s}, \cdots \frac{1}{A_r^s}, \cdots -\frac{A_m^s}{A_r^s} \right]$$

This is sufficient information to reconstruct the pivot matrix.

Product Form of the Inverse

If matrix A is nonsingular, then a sequence of pivots down the diagonal of A (with possible row interchanges to avoid zero pivot elements) will reduce A to the identity matrix. This is equivalent to pre-multiplying A by a sequence of pivot matrices:

$$(P_m \cdots (P_3(P_2(P_1 A))) \cdots) = I$$

$$\Longrightarrow \quad (P_m \cdots P_3 P_2 P_1) A = I$$

$$\Longrightarrow \quad A^{-1} = P_m \cdots P_3 P_2 P_1$$

Product Form of the Inverse

In the Revised Simplex Method, computation of values in the tableau is done, not by pivoting in the tableau, but by either pre-multiplication or post-multiplication by the inverse matrix:

- Computation of simplex multipliers

$$\pi = c^B (A^B)^{-1}$$

used in selecting pivot column

- Computation of substitution rates

$$\alpha = (A^B)^{-1} A^s$$

used in performing the pivot

Computing Simplex Multipliers

Solve $\pi A^B = c^B$ for π :

$$\pi = c^B (A^B)^{-1}$$
$$= c^B (P_k P_{k-1} \cdots P_3 P_2 P_1)$$
$$= (((\cdots(c^B P_k) P_{k-1} \cdots P_3) P_2) P_1)$$

"Backward Transformation", or BTRAN

The pivot matrices are processed in the ***reverse*** of the order in which they were generated, i.e., $P_k P_{k-1} \ldots P_3 P_2 P_1$

BTRAN

For each pivot matrix P, we need to calculate $\pi = \nu P$

column r

$$\pi = [\,\nu_1 \;\; \nu_2 \;\; \cdots \;\; \nu_{m-1} \;\; \nu_m\,] \begin{bmatrix} 1 & 0 \cdots & \eta_1 & \cdots 0 & 0 \\ 0 & 1 \cdots & \eta_2 & \cdots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 \cdots & \eta_r & \cdots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 \cdots & \eta_{m-1} & \cdots 1 & 0 \\ 0 & 0 \cdots & \eta_m & \cdots 0 & 1 \end{bmatrix}$$

$$= \left[\,\nu_1 \;\; \nu_2 \;\; \cdots \;\; \left(\sum_i \nu_i \eta_i\right) \;\; \cdots \;\; \nu_{m-1} \;\; \nu_m\,\right]$$

entry r

BTRAN

$$\pi_j = \begin{cases} \nu_j & \text{for } j \neq r \\ \sum_i \nu_i \eta_i & \text{for } j = r \end{cases}$$

Step 0: Set $\nu = c^B$ and k = # of ETA vectors

Step 1: Using BTRAN formula above, compute with ETA vector #k

Step 2: If k>1, let $\nu = \pi$ and k = k−1, and go to step 1; else proceed to step 3.

Step 3: The final value of π is the solution of $\pi A^B = c^B$

FTRAN

Solve $A^B \alpha = A^s$ for substitution rates α

$$\alpha = (A^B)^{-1} A^s$$
$$= (P_k P_{k-1} \cdots P_3 P_2 P_1) A^s$$
$$= (P_k (P_{k-1} \cdots P_3 (P_2 (P_1 A^s)) \cdots))$$

"Forward Transformation", or FTRAN

The pivot matrices are processed in the same order that they were generated, i.e., $P_1, P_2, P_3, \cdots \quad P_{k-1}, P_k$

FTRAN

column r

$$\alpha = \begin{bmatrix} 1 & 0 & \cdots & \eta_1 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & \eta_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \eta_r & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \eta_{m-1} & \cdots & 1 & 0 \\ 0 & 0 & \cdots & \eta_m & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_m \end{bmatrix} = \begin{bmatrix} \nu_1 + \eta_1 \nu_r \\ \nu_2 + \eta_2 \nu_r \\ \vdots \\ \eta_r \nu_r \\ \nu_m + \eta_m \nu_r \end{bmatrix}$$

That is,

$$\alpha_i = \begin{cases} \nu_i + \eta_i \nu_r & \text{for } i \neq r \\ \eta_r \nu_r & \text{for } i = r \end{cases}$$

FTRAN

$$\alpha_i = \begin{cases} \nu_i + \eta_i \nu_r & \text{for } i \neq r \\ \eta_r \nu_r & \text{for } i = r \end{cases}$$

Step 0: Set $\nu = A^s$ (e.g., column of original tableau), and k=1.

Step 1: Using the FTRAN formula above, compute α

Step 2: If k < # of ETA vectors, then let $\nu = \alpha$ and k=k+1, and go to step 1; else proceed to step 3.

Step 3: The final value of ν is the solution α of the equation $A^B \alpha = A^s$

Gauss Elimination

-- a method for solving Ax=b by performing a sequence of elementary row operations on the augmented matrix $[A \mid b]$ to reduce it to an echelon matrix. The solution is then obtained by "back-substitution".

Example:

$$\begin{cases} x_1 + x_2 + x_3 = 4 \\ x_1 + 2x_2 + 2x_3 = 2 \\ -x_1 - x_2 + x_3 = 2 \end{cases}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 1 & 2 & 2 & 2 \\ -1 & -1 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right] \Longrightarrow \begin{cases} x_1 + x_2 + x_3 = 4 \\ x_2 + x_3 = -2 \\ x_3 = 3 \end{cases}$$

Backsubstitution:

$$\left\{ \begin{array}{l} x_1 = 4 - x_2 - x_3 \\ \left. \begin{array}{l} x_2 = -2 - x_3 \\ x_3 = 3 \end{array} \right\} \Longrightarrow x_2 = -5 \end{array} \right\} \Longrightarrow x_1 = 6$$

Gauss-Jordan Elimination

--similar to Gauss elimination, except that the coefficient matrix is diagonalized by further elementary row operations, eliminating non-zeroes above as well as below the diagonal. Eliminates the need for "back-substitution".

Example:

$$\begin{cases} x_1 + x_2 + x_3 = 4 \\ x_1 + 2x_2 + 2x_3 = 2 \\ -x_1 - x_2 + x_3 = 2 \end{cases}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 1 & 2 & 2 & 2 \\ -1 & -1 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

That is,

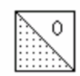
$$\begin{cases} x_1 = 6 \\ x_2 = -5 \\ x_3 = 3 \end{cases}$$

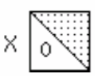
Compared to "Gauss Elimination Plus Back Substitution", Gauss-Jordan Elimination requires more computation-- especially if the equations are to be solved for several right-hand-side vectors!

Gauss Elimination as Matrix Factorization

A = P L U

P is a permutation matrix (which performs the interchange of rows for partial pivoting)

L is a lower triangular matrix, 

U is an upper triangular matrix 

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

$R_2 \leftarrow R_2 + R_1$ $\quad$ $R_3 \leftarrow R_3 - R_2$

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Upper-triangular matrix

Lower triangular matrices

$$\underbrace{E_2 E_1}_{\hat{L}} A = U$$

Lower triangular matrix

$$\hat{L} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}, \quad \hat{L}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = L$$

$$\hat{L}\, A = U \implies A = \hat{L}^{-1} U = L\, U$$

Matrix A is factored into a product of lower & upper triangular matrices!

Suppose that we need to solve

$$\begin{cases} x_1 + 2x_2 + x_3 = 2 \\ -x_1 - x_2 + x_3 = 5 \\ x_2 + 3x_3 = -1 \end{cases}$$

$$A = L\,U = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

To solve Ax=b, i.e., L(Ux)=b:

- solve Ly=b for y *(forward substitution)*

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} y = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} \implies \begin{cases} y_1 = 2 \\ y_2 = 5 + y_1 = 7 \\ y_3 = -1 - y_2 = -8 \end{cases}$$

- solve Ux=y for x *(backward substitution)*

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 2 \\ 7 \\ -8 \end{bmatrix} \implies \begin{cases} x_1 = 2 - 2x_2 - x_3 = -36 \\ x_2 = 7 - 2x_3 = 23 \\ x_3 = -8 \end{cases}$$

CHOLESKY FACTORIZATION

Suppose that A is a ***symmetric*** & ***positive definite*** matrix.

Then the ***Cholesky factorization*** of A is

$$A = \hat{L}\,\hat{L}^T$$

where $\hat{L}$ is a ***lower triangular*** matrix.

Computation:
Suppose that we have the factorization

$$A = L\,D\,L^T$$

Then if $D_i^i \ge 0$, we can define a new diagonal matrix $\hat{D}$ where

$$\hat{D}_i^i \equiv \sqrt{D_i^i}$$

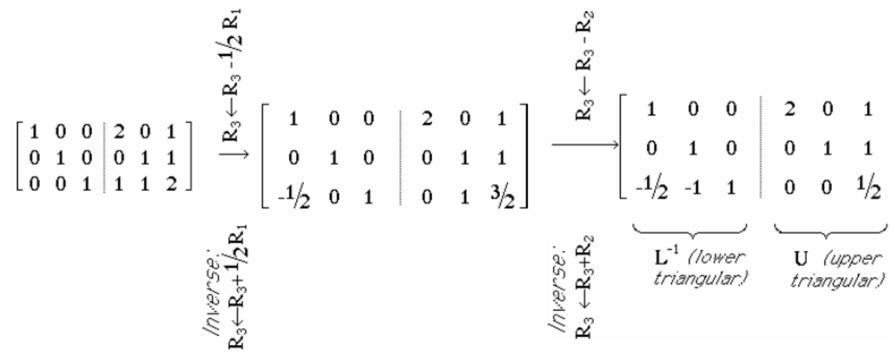
Then $A = L\,D\,L^T = L\,\hat{D}\,\hat{D}\,L^T = \left(L\hat{D}\right)\,\left(L\hat{D}\right)^T = \hat{L}\,\hat{L}^T$ where $\hat{L} = L\hat{D}$

Example:

We wish to find the Cholesky factorization of the matrix

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

Cholesky factorization…



The lower triangular matrix L is found by performing (on the identity matrix) the inverse of the row operations used to reduce the A matrix:

$$\left.\begin{array}{l} R_3 \leftarrow R_3 + \tfrac{1}{2} R_1 \\ R_3 \leftarrow R_3 + R_2 \end{array}\right\} \Rightarrow L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \tfrac{1}{2} & 1 & 1 \end{bmatrix}$$

We now have the LU factorization of matrix A:

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \tfrac{1}{2} & 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & \tfrac{1}{2} \end{bmatrix}$$

Define the diagonal matrix D:

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \tfrac{1}{2} \end{bmatrix} \Rightarrow D^{-1} = \begin{bmatrix} \tfrac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Note that

$$\hat{U} = D^{-1}U = \begin{bmatrix} \tfrac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \tfrac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \tfrac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

And so,

$$A = LDL^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \tfrac{1}{2} & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \tfrac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & \tfrac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Define the diagonal matrix $\hat{D}$ where $\hat{D}_i^i \equiv \sqrt{D_i^i}$:

$$\hat{D} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \tfrac{1}{\sqrt{2}} \end{bmatrix}$$

Then compute $\hat{L} = L\hat{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \tfrac{1}{2} & 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \tfrac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ \tfrac{1}{\sqrt{2}} & 1 & \tfrac{1}{\sqrt{2}} \end{bmatrix}$

So the Cholesky factorization is

$$A = \hat{L}\hat{L}^T = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ \tfrac{1}{\sqrt{2}} & 1 & \tfrac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & \tfrac{1}{\sqrt{2}} \\ 0 & 1 & 1 \\ 0 & 0 & \tfrac{1}{\sqrt{2}} \end{bmatrix}$$