In [1]:
```
import sympy
from sympy import Matrix, Rational, sqrt, symbols
import numpy as np
%matplotlib widget
import matplotlib.pyplot as plt
```

Symmetric matrices

- symmetric matrices are square matrices S with the property that

$$S = S^T$$

- if there is an LDU decomposition for a symmetric matrix S, then

$$S = LDU$$
$$L = U^T$$

- in other words, a symmetric matrix S can be decomposed as

$$S = LDL^T$$

- if row permutation is required, it has to be accompanied by column permutation to preserve symmetry

example

In [2]:
```
S = Matrix([
    [0, 1, 2],
    [1, -1, 1],
    [2, 1, 3]
])
S
```

Out[2]:
$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & -1 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

- permutation matrix

In [3]:
```
l, u, p = S.LUdecomposition()
```

In [4]:
```
l
```

Out[4]:
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix}$$

In [5]:
```
u
```

Out[5]:

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

In [6]:
```
p
```

Out[6]:
```
[[0, 1]]
```

In [7]:
```
l * u
```

Out[7]:

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

In [8]:
```
P = Matrix([
    [0,0,1],
    [0,1,0],
    [1,0,0]
])
P
```

Out[8]:

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

In [9]:
```
S1 = P * S * P.T
S1
```

Out[9]:

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

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

$$D = \begin{bmatrix}3 & 0 & 0\\0 & -\frac{4}{3} & 0\\0 & 0 & -\frac{5}{4}\end{bmatrix}$$

$$PSP^T = LDL^T$$

Cholesky decomposition

Let us consider the so-called **quadratic form** of S:

$$\mathbf{x}' S \mathbf{x}$$

This function assigns each vector $\mathbf{x}$ to some real number – something like a high-dimensional parabola.

Example:

In [10]:

```
S = Matrix([
    [3, 1],
    [1, -1]
])
S
```

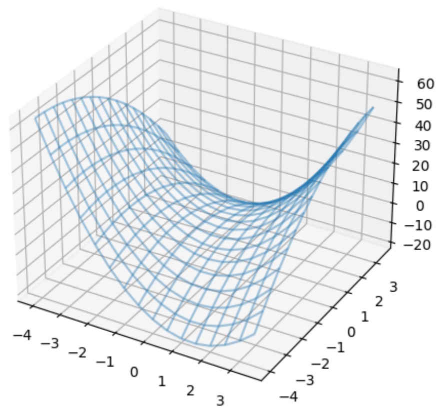
Out[10]:

$$\begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix}$$

In [11]:

```
fig = plt.figure(figsize=(12,6))
ax = fig.add_subplot(121, projection='3d')

x = np.arange(-4, 4, .5)
y = np.arange(-4, 4, .5)
X,Y = np.meshgrid(x,y)


def sQuadratic(v, S=S):
    return (S @ v) @ v

Z = np.apply_along_axis(sQuadratic, 0, np.array([X, Y]))

ax.plot_wireframe(X, Y, Z, alpha=0.4)
plt.show()
```

Figure



Positive definite matrices

The quadratic form of S has a saddle point at $\mathbf{0}$.

Symmetric matrices with a quadratic form which is everywhere $\neq \mathbf{0}$ positive are called **positive definit**.

Definition

A matrix S is **positive definite** if and only if for all vectors $\mathbf{x} \neq \mathbf{0}$:

$$\mathbf{x}'S\mathbf{x} > 0$$

Example

In [12]:

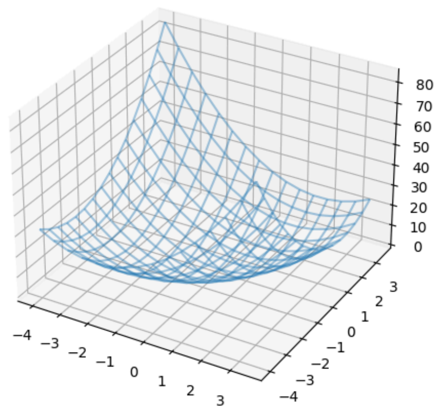
```
S1 = Matrix([
    [2,-1],
    [-1,2]
])
S1
```

Out[12]: $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

In [13]:

```
fig = plt.figure(figsize=(12,6))
ax = fig.add_subplot(121, projection='3d')


Z = np.apply_along_axis(lambda v:sQuadratic(v, S1), 0, np.array([X, Y]))

ax.plot_wireframe(X, Y, Z, alpha=0.4)
plt.show()
```

Figure



Suppose we know that

$$\begin{aligned} \mathbf{x}^T S \mathbf{x} > 0 & \qquad \forall \mathbf{x} : \mathbf{x} \neq \mathbf{0} \\ S = LDL^T & \end{aligned}$$

This entails

$$\begin{aligned} \mathbf{x}^T S \mathbf{x} > 0 \\ \mathbf{x}^T LDL^T \mathbf{x} > 0 \end{aligned}$$

Let us introduce a variable $\mathbf{y}$

$$\begin{aligned} \mathbf{y} &= L^T \mathbf{x} \\ \mathbf{x} &= (L^T)^{-1} \mathbf{y} \qquad (L \text{ must be invertible, which follows from definiteness.}) \\ \mathbf{y}^T D \mathbf{y} &> 0 \\ \sum_i d_i y_i^2 &> 0 \end{aligned}$$

For each i, we can set $y_i = 1, \forall j \neq i : y_j = 0$. It follows that

$$\forall i : d_i > 0$$

So if S is positive definite and the LU-decomposition gives $S = LDL^T$, all entries of D must be positive.

We define

$$(D^{\frac{1}{2}})_{ij} = \sqrt{d_{ij}}$$

It follows that

$$D = D^{\frac{1}{2}} D^{\frac{1}{2}}$$

Taken together, we have

$$S = LD^{\frac{1}{2}}D^{\frac{1}{2}}L^T$$

It follows that for each positive definite matrix S wich has an LU decomposition, there is a lower triangular matrix M such that

$$S = MM^T$$

This factorization is called **Cholesky decomposition**.

In [14]:
```
S = Matrix([
    [2,-1,0],
    [-1,2,-1],
    [0,-1,2]
])
S
```

Out[14]:

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

In [15]:
```
S.is_positive_definite
```

Out[15]:
```
True
```

In [16]:
```
M = S.cholesky()
M
```

Out[16]:

$$\begin{bmatrix} \sqrt{2} & 0 & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{2} & 0 \\ 0 & -\frac{\sqrt{6}}{3} & \frac{2\sqrt{3}}{3} \end{bmatrix}$$

In [17]:
```
M*M.T
```

Out[17]:

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

In [18]:
```
S = Matrix([
    [3, 1],
    [1, -1]
])
S
```

Out[18]: $\begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix}$

In [19]:
```
S.is_positive_definite
```

Out[19]:
```
False
```

In [20]:
```
S.cholesky()
```

```
---------------------------------------------------------------------------
NonPositiveDefiniteMatrixError            Traceback (most recent call last)
/tmp/ipykernel_45595/1885332215.py in <cell line: 1>()
----> 1 S.cholesky()

~/miniconda3/envs/math_ml_24/lib/python3.9/site-packages/sympy/matrices/dense.py
in cholesky(self, hermitian)
     79
     80     def cholesky(self, hermitian=True):
---> 81         return _cholesky(self, hermitian=hermitian)
     82
     83     def LDLdecomposition(self, hermitian=True):

~/miniconda3/envs/math_ml_24/lib/python3.9/site-packages/sympy/matrices/decompos
itions.py in _cholesky(M, hermitian)
    272
    273             if Lii2.is_positive is False:
--> 274                 raise NonPositiveDefiniteMatrixError(
    275                     "Matrix must be positive-definite")
    276

NonPositiveDefiniteMatrixError: Matrix must be positive-definite
```

Fact

If a square matrix (not necessarily symmetric) is invertible, the following two matrices are positive definite:

$$AA^T$$
$$A^TA$$

Mathematics for Machine Learning

Session 05:

Gerhard Jäger

November 5, 2024

Vector spaces

In mathematics, vector spaces are not confined to $\mathbb{R}^n$. Any set can be a vector space if the following eight axioms are met.

Axiom	Meaning
Associativity of vector addition	**u** + (**v** + **w**) = (**u** + **v**) + **w**
Commutativity of vector addition	**u** + **v** = **v** + **u**
Identity element of vector addition	There exists an element **0** ∈ *V*, called the *zero vector*, such that **v** + **0** = **v** for all **v** ∈ *V*.
Inverse elements of vector addition	For every **v** ∈ *V*, there exists an element −**v** ∈ *V*, called the *additive inverse* of **v**, such that **v** + (−**v**) = **0**.
Compatibility of scalar multiplication with field multiplication	*a*(*b***v**) = (*ab*)**v** [nb 3]
Identity element of scalar multiplication	1**v** = **v**, where 1 denotes the multiplicative identity in $\mathbb{R}$.
Distributivityof scalar multiplication with respect to vector addition	*a*(**u** + **v**) = *a***u** + *a***v**
Distributivity of scalar multiplication with respect to field addition	(*a* + *b*)**v** = *a***v** + *b***v**

A subset α of a vector space is a **sub-space** if it is also a vector space.

This is guaranteed if two conditions are met: If $\mathbf{u}, \mathbf{v} \in \alpha$,

- $\mathbf{u} + \mathbf{v} \in \alpha$
- for any scalar c, $c\mathbf{u} \in \alpha$.

Column space of a matrix

For an $n \times m$ matrix A, the set $\{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^m\}$ is a subspace of $\mathbb{R}^n$, because:

- $A\mathbf{u} + A\mathbf{v} = A(\mathbf{u} + \mathbf{v})$
- $Ac\mathbf{u} = cA\mathbf{u}$

This vector space is called the **column space** of A, because it consists of all linear combinations of the column vectors of A.

The system of linear equation

$$A\mathbf{x} = \mathbf{b}$$

is solvable if and only if $\mathbf{b}$ is in the column space of A.

So far we have focused on invertable matrices. In this case, the column space equals the entire vector space.

But what about non-invertable matrices?

We proceed per Gauss-Jordan elimination as usual. After the Gauss part, we will see whether the system is solvable.

$$\begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 4 & 3 & 1 \\ 2 & 3 & 0 \end{array}\right]$$

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

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

There can't be a solution for the last row, so $\mathbf{b}$ is not in the column space of A.

Another example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 3 \\ 3 & 6 & 1 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 4 \\ 5 \\ 4 \end{bmatrix}$$

$$A\mathbf{x} = \mathbf{b}$$

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

$$\left[\begin{array}{ccr|r} 1 & 2 & 3 & 4 \\ 0 & 0 & -3 & -3 \\ 3 & 6 & 1 & 4 \end{array}\right]$$

$$\left[\begin{array}{ccr|r} 1 & 2 & 3 & 4 \\ 0 & 0 & -3 & -3 \\ 0 & 0 & -8 & -8 \end{array}\right]$$

$$\left[\begin{array}{ccr|r} 1 & 2 & 3 & 4 \\ 0 & 0 & -3 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccr|r} 1 & 2 & 0 & 1 \\ 0 & 0 & -3 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Since the left side is not the identity matrix, we cannot simply read off the solution from the right side.

Pivots

A pivot cell in a matrix A is a cell a_{ij} with an entry $\neq 0$ such that

$$\begin{aligned} a_{ij} &\neq 0 \\ \forall k \geq i, l < j : a_{kl} &= 0 \end{aligned}$$

During Gauss elimination, we

- go top-down through the rows,
- identify the pivot in the current row, and
- transform all cells below the pivot into 0 using elimination.

A column with a pivot in it is called *pivot column*.

A column without a pivot in it is called *free column*.

After Gauss-Jordan elimination, we find the *canonical solution* by

- setting all positions corresponding to a free column to 0, and
- setting all positions corresponding to a pivot colun to the entry on the the same row as the pivot on the right side.

For our example above, the solution thus found is

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

The nullspace of a matrix

The canonical solution is only one of infinitely many solutions.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 3 \\ 3 & 6 & 1 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 4 \\ 5 \\ 4 \end{bmatrix}$$

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}, \cdots \end{aligned}$$

How do we find all solutions?

Suppose we have a solution $\mathbf{y}$ to

$$A\mathbf{y} = \mathbf{0}$$

Let $\mathbf{x}$ be the canonical solution to

$$A\mathbf{x} = \mathbf{b}$$

It follows:

$$\begin{aligned} A\mathbf{x} + A\mathbf{y} &= \mathbf{b} + \mathbf{0} \\ A(\mathbf{x} + \mathbf{y}) &= \mathbf{b} \end{aligned}$$

The set of solutions $\mathbf{y}$ to $A\mathbf{y} = \mathbf{0}$ form a *vector space*:

Suppose $A\mathbf{y}_1 = \mathbf{0}$ and $A\mathbf{y}_2 = \mathbf{0}$.

$$\begin{aligned} A(\mathbf{y}_1 + \mathbf{y}_2) &= A\mathbf{y}_1 + A\mathbf{y}_2 \\ &= \mathbf{0} + \mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

$$\begin{aligned} A(c\mathbf{y}_1) &= cA\mathbf{y}_1 \\ &= c\,\mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

The set of solutions $\mathbf{y}$ to $A\mathbf{y} = \mathbf{0}$ is called the **nullspace** of A.

How to find the null space

important observation: applying an elimination step to a matrix does not change the null space!

If

$$A\mathbf{x} = \mathbf{0},$$

then

$$EA\mathbf{x} = E\mathbf{0} = \mathbf{0}$$

For the reverse direction, note that elimination matrices are always invertible by construction.

If

$$EA\mathbf{x} = \mathbf{0}$$

then

$$A\mathbf{x} = E^{-1}\mathbf{0} = \mathbf{0}$$

So we can apply Gauss-Jordan elimination first and then find the null space of the **reduced row echelon form**.

For each free column i in the reduced row echelon form or A:

- set $y_i = 1$,
- set $y_j = 0$ for all free columns $j \neq i$,
- solve $A\mathbf{y} = \mathbf{0}$ via substitution.

1 The **nullspace** $\boldsymbol{N}(A)$ in $\mathbf{R}^n$ contains all solutions $\boldsymbol{x}$ to $A\boldsymbol{x} = \mathbf{0}$. This includes $\boldsymbol{x} = \mathbf{0}$.

2 Elimination (from A to U to R) does not change the nullspace: $\boldsymbol{N}(A) = \boldsymbol{N}(U) = \boldsymbol{N}(R)$.

3 The **reduced row echelon form** $\boldsymbol{R} = \mathbf{rref}(\boldsymbol{A})$ has all pivots $= 1$, with zeros above and below.

4 If column j of R is free (no pivot), there is a "*special solution*" to $A\boldsymbol{x} = \mathbf{0}$ with $x_j = 1$.

5 Number of pivots = number of nonzero rows in R = **rank** $\boldsymbol{r}$. There are $n - r$ free columns.

6 Every matrix with $m < n$ has nonzero solutions to $A\boldsymbol{x} = \mathbf{0}$ in its nullspace.

(Strang, p. 135)

simple example

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$$

Elimination stops after one step:

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

- the entry in the upper left corner is the only pivot
- the second column is free
- the special solution is the solution of

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ 1 \end{bmatrix} = \mathbf{0}$$

- the only special solution is $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$
- the null space of A is the set of multiples of $\mathbf{x}$

The result of Gauss elimination is called **row echelon form**.

The result of Gauss-Jordan elimination is called **reduced row echelon form**.

Sympy has a function that returns

- the reduced row echelon form of a matrix
- the tuple of indices of the pivot columns

In [21]:

```
A = Matrix([
    [1,2],
    [3,6]
])
rr,i = A.rref()
rr
```

Out[21]: $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$

In [22]:

```
i
```

Out[22]:

```
(0,)
```

back to our previous examples

In [23]:

```
A = Matrix([
    [1,2,3],
    [2,4,3],
    [3,6,1]
])
A
```

Out[23]: $\begin{bmatrix}1 & 2 & 3\\2 & 4 & 3\\3 & 6 & 1\end{bmatrix}$

In [24]:
```
rr,i = A.rref()
rr
```

Out[24]: $\begin{bmatrix}1 & 2 & 0\\0 & 0 & 1\\0 & 0 & 0\end{bmatrix}$

In [25]:
```
i
```

Out[25]:
```
(0, 2)
```

$$\begin{bmatrix}1 & 2 & 0\\0 & 0 & 1\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}a\\1\\b\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$
$$a + 2 = 0$$
$$b = 0$$
$$a = -2$$
$$\mathbf{y} = \begin{bmatrix}-2\\1\\0\end{bmatrix}$$

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

example

$$A = \begin{bmatrix}1 & 2\\3 & 8\end{bmatrix}$$
$$A\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The last matrix contains the *reduced row echelon form* of A (rref(A). It contains two pivots and now free columns.

Hence the only solution is $\mathbf{x} = \mathbf{0}$.

example

$$B = \begin{bmatrix} A \\ 2A \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 4 \\ 6 & 16 \end{bmatrix}$$
$$B\mathbf{y} = \mathbf{0}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 4 \\ 6 & 16 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 2 & 4 \\ 6 & 16 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 0 & 0 \\ 6 & 16 \end{bmatrix}$$

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

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

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

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

rref(B) contains no free columns.

Again, the only solution is $\mathbf{0}$

example

$$\begin{aligned} C &= \begin{bmatrix} A & 2A \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & 2 & 4 \\ 3 & 8 & 6 & 16 \end{bmatrix} \\ C\mathbf{z} &= 0 \end{aligned}$$

$$\begin{bmatrix} 1 & 2 & 2 & 4 \\ 3 & 8 & 6 & 16 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 2 & 0 & 4 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 4 \end{bmatrix}$$
$$\begin{bmatrix} \mathbf{1} & 0 & 2 & 0 \\ 0 & \mathbf{1} & 0 & 2 \end{bmatrix}$$

The last matrix contains rref(C). This time we have two free columns, the third and the fourth.

- special solutions:

$$\mathbf{z}_1 = \begin{bmatrix} a \\ b \\ 1 \\ 0 \end{bmatrix}$$
$$\mathbf{z}_2 = \begin{bmatrix} c \\ d \\ 0 \\ 1 \end{bmatrix}$$

- solving via substitution

$$\begin{aligned} \operatorname{rref}(C)\ \mathbf{z}_1 &= \mathbf{0} \\ \begin{bmatrix} \mathbf{1} & 0 & 2 & 0 \\ 0 & \mathbf{1} & 0 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ 1 \\ 0 \end{bmatrix} &= \mathbf{0} \\ a + 2 &= 0 \\ b &= 0 \\ a &= -2 \\ \mathbf{z}_1 &= \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

- solving via substitution

$$\mathrm{rref}(C)\ \mathbf{z}_2 = \mathbf{0}$$

$$\begin{bmatrix} \mathbf{1} & 0 & 2 & 0 \\ 0 & \mathbf{1} & 0 & 2 \end{bmatrix} \begin{bmatrix} c \\ d \\ 0 \\ 1 \end{bmatrix} = \mathbf{0}$$

$$c = 0$$

$$d + 2 = 0$$

$$d = -2$$

$$\mathbf{z}_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

Any linear combination

$$\mathbf{z} = \alpha \mathbf{z}_1 + \beta \mathbf{z}_2$$

(α, β real numbers) are solutions to

$$C\mathbf{z} = \mathbf{0}$$

The complete solution to a system of equations

Back to our example from above:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 3 \\ 3 & 6 & 1 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 4 \\ 5 \\ 4 \end{bmatrix}$$

$$A\mathbf{x} = \mathbf{b}$$

- first step: apply Gauss-Jordan elimination to the augmented matrix

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

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & -3 & -3 \\ 0 & 0 & -8 & -8 \end{array}\right]$$

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

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

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

- second step: find canonical solution by setting the slots corresponding to the free columns to 0 and solving via substitution

$$\mathbf{x}_c = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

- third step: find the null space of A

$$\mathbf{x}_0 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

The complete solution is the set of vectors

$$\mathbf{x}_c + \alpha\, \mathbf{x}_0,$$

for any real number α.

Rank of a matrix

Geometrically speaking, the **rank** of a matrix is the number of dimensions covered by its column space.

examples of an invertable and a non-invertable matrix in 3d

$$A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 1 & 3 \\ 2 & 6 & 8 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} \frac{5}{63} & -\frac{22}{63} & \frac{1}{9} \\ -\frac{11}{63} & -\frac{2}{63} & \frac{1}{18} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{18} \end{bmatrix}$$

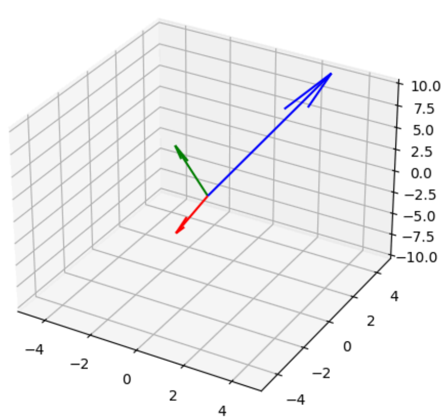
$$B = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 1 & 3 \\ 2 & 6 & -10 \end{bmatrix}$$

B^{-1} is undefined

In [26]:

```
fig = plt.figure(figsize=(12,6))
ax = fig.add_subplot(121, projection='3d')



ax.set_xlim(-5,5)
ax.set_ylim(-5,5)
ax.set_zlim(-10,10)
ax.quiver((0,),(0,),(0,),(1,),(-4,),(2,), color='red', length=1)
ax.quiver((0,),(0,),(0,),(-2,),(1,),(3,), color='green', length=1)
ax.quiver((0,),(0,),(0,),(2,),(6,),(8,), color='blue', length=1)

plt.show()
```

Figure

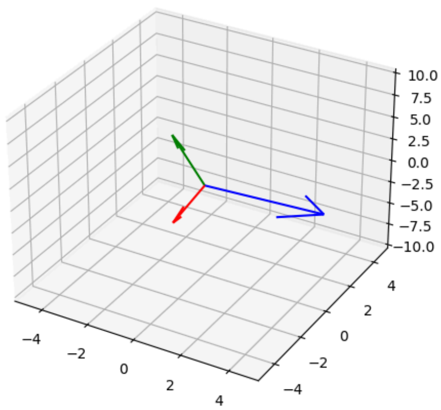


In [27]:

```
fig = plt.figure(figsize=(12,6))
ax = fig.add_subplot(121, projection='3d')
```

```
ax.set_xlim(-5,5)
ax.set_ylim(-5,5)
ax.set_zlim(-10,10)
ax.quiver((0,),(0,),(0,),(1,),(-4,),(2,), color='red', length=1)
ax.quiver((0,),(0,),(0,),(-2,),(1,),(3,), color='green', length=1)
ax.quiver((0,),(0,),(0,),(2,),(6,),(-10,), color='blue', length=1)

plt.show()
```

Figure



- the invertable matrix has rank 3
- the non-invertable matrix has rank 2

Fact:

The rank of a matrix equals the number of pivots in its row echelon form, which is equal to the number of pivots in its reduced row echelon form.

- rank-3 matrix:

$$A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 1 & 3 \\ 2 & 6 & 8 \end{bmatrix}$$

- row echelon form:

$$\begin{bmatrix} 1 & -4 & 2 \\ -2 & 1 & 3 \\ 2 & 6 & 8 \end{bmatrix}$$
$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & -7 & 7 \\ 0 & 14 & 4 \end{bmatrix}$$
$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & -7 & 7 \\ 0 & 0 & 28 \end{bmatrix}$$

The row echelon form has 3 pivots, so A has rank 3.

$$\begin{bmatrix} 1 & -4 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 28 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 28 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The reduced row echolon form also has 3 pivots. (For an invertable matrix, the reduced row echelon form is always the identity matrix.)

- rank-2 matrix

$$B = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 1 & 3 \\ 2 & 6 & -10 \end{bmatrix}$$

- row-echelon form

$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & -7 & 7 \\ 0 & 14 & -14 \end{bmatrix}$$
$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & -7 & 7 \\ 0 & 0 & 0 \end{bmatrix}$$

The row echelon form has two pivots, hence B has rank 2.

- reduced row echelon form

$$\begin{bmatrix}1 & 0 & -2\\0 & -7 & 7\\0 & 0 & 0\end{bmatrix}$$

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

The reduced row echelon form also has 2 pivots.

Dimensionality of a vector space

Linear independence of vectors

- a vector $\mathbf{b}$ is independent from a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ if and only if there are no real numbers $x_1, \ldots, x_n$ such that

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

- a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is linearly independent if and only if each $\mathbf{a}_i$ is linearly independent of the other vectors $\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots \mathbf{a}_n$.

Equivalently, the column vectors of a matrix A are linearly independent if and only if the linear system

$$A\mathbf{x} = \mathbf{0}$$

has as its only solution $\mathbf{x} = \mathbf{0}$.

(Independence is important, e.g., for linear regression.)

Examples

- $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$: dependent or independent?
- $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\0.00001\end{bmatrix}$: dependent or independent?
- $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}-1\\-1\end{bmatrix}$: dependent or independent?
- $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}0\\0\end{bmatrix}$: dependent or independent?
- $\begin{bmatrix}a\\b\end{bmatrix}$, $\begin{bmatrix}c\\d\end{bmatrix}$ and $\begin{bmatrix}e\\f\end{bmatrix}$: dependent or independent?

In [28]:

```
A = Matrix([
    [1,0,3],
    [2,1,5],
```

```
    [1,0,3]
])
A
```

Out[28]:
$$\begin{bmatrix}1 & 0 & 3\\2 & 1 & 5\\1 & 0 & 3\end{bmatrix}$$

- are the columns dependent or independent?
- what is the rank of A?

In [29]:
```
A.nullspace()[0]
```

Out[29]:
$$\begin{bmatrix}-3\\1\\1\end{bmatrix}$$

A little puzzle

A rectangular $m \times n$ matrix with $m < n$ cannot have independent columns. Why?

Example

In [30]:
```
B = Matrix([
    [1,0,1,5],
    [0,1,1,-3],
    [2,1,-1,-2]
])
B
```

Out[30]:
$$\begin{bmatrix}1 & 0 & 1 & 5\\0 & 1 & 1 & -3\\2 & 1 & -1 & -2\end{bmatrix}$$

In [31]:
```
B.rref()[0]
```

Out[31]:
$$\begin{bmatrix}1 & 0 & 0 & \frac{11}{4}\\0 & 1 & 0 & -\frac{21}{4}\\0 & 0 & 1 & \frac{9}{4}\end{bmatrix}$$

Basis of a vector space

- the **span** of a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is the set of their linear combinations, i.e., the set

$$\{x_1\mathbf{a}_1 + \cdots x_n\mathbf{a}_n | x_1, \ldots, x_n \in \mathbb{R}\}$$

- the span of a set of vectors is a vector space

- a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is a **basis of a vector space** S if and only if
  - $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent, and
  - the span of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is S.

What is the span of

- $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,
- $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$,
- $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$?

Which of these sets of vectors form a basis?