MAT 140

Linear Algebra and Affine Geometry

Dr Ton Boerkoel

DigiPen Institute of Technology

Fall 2017

Contents

Introduction

Part 1

An Exploration of 2 and 3 Dimensional Euclidean Space

1. Points and Vectors	1
2. The Dot Product	9
3. A Basic Projection	13
4. Vector Equations of Lines	19
5. The Distance of a Point to a Line in 2D	23
6. Lines and Planes in 3D	27
7. The Cross Product	33
8. Intersections	39
9. Angles	45
10. Distances	49

Part II

Matrices, Systems of Equations and Row reduction

11. Matrices	59
12. Determinants	71
13. Properties of Determinants	81
14. Reduced Row Echelon Form	85
15. Solving Systems of Equations	97

Part III

Linear Transformations

16. Linear Transformations	113
17. Compositions of Linear Transformations	127
18. Basic Linear Transformation Matrices in 2D and 3D	131
19. 2D projections	135
20. 2D reflections	147
21. 2D scaling	153
22. 2D rotations	155
23. 2D shears	161

24. 3D projections 167
25. 3D reflections 183
26. 3D scaling 193
27. 3D shears 195
28. 3D rotations 201
29. Traces, Determinants and Eigenvalues 215

Part IV

Affine Transformations

30. Translation 227
31. Affine Transformations 231
32. More on Affine Transformations 237
33. Basic Affine Transformations 247
34. Barycentric Coordinates 257
35. More on Barycentric Coordinates 273
36. An Extra Dimension 287
37. The Extra Dimension at work 295
38. The Main Affine Matrices and Homogeneous Coordinates 309
39. Perspective Projection 327

11. Matrices

A matrix is a rectangular array of numbers. In this course we will only deal with matrices that have real numbers as entries. For example $\begin{bmatrix} 3 & 4 & -1 \\ 2 & -7 & 0 \end{bmatrix}$ is a 2×3 matrix with real entries, in fact in this case all its entries are integers. Here is another 2×3 matrix with real entries.

$$\begin{bmatrix} \sqrt{3} & 3+\ln(4) & \pi \\ e^3-4 & 4/7 & \sin(2.5) \end{bmatrix}$$

In general an $n\times m$ matrix A looks like

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

where all the entries a_{ij} are real numbers. Sometimes we write $A = \left[a_{ij} \right]_{\substack{i=1,2,\cdots,n \\ j=1,2,\cdots,m}}$.

Note that an $n\times m$ matrix has n rows and m columns:

$$\begin{array}{cc} & \longleftarrow \; m \; \longrightarrow \\ \updownarrow n & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \end{array}$$

The entry in row i and column j is indicated by a_{ij}. The set of all $n\times m$ matrices with real entries we will denote by $M_{n\times m}(\mathbb{R})$. Note that a vector $\begin{bmatrix} x \\ y \end{bmatrix}$ can be seen as 2×1 matrix, i.e. $\begin{bmatrix} x \\ y \end{bmatrix} \in M_{2\times 1}(\mathbb{R})$; and $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in M_{3\times 1}(\mathbb{R})$.

Matrix operations

(1) Matrix addition

We will *only* add matrices with the *same* dimensions (same number of rows and same number of columns). Addition is performed component wise, e.g.

$$\begin{bmatrix} 3 & 4 & -1 \\ 2 & -7 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -2 & 8 \\ 3 & 4 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 7 \\ 5 & -3 & -1 \end{bmatrix}$$

In general we have

$$A+B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2m}+b_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nm}+b_{nm} \end{bmatrix}$$

or more condensed $A+B = \left[a_{ij}\right]+\left[b_{ij}\right]=\left[a_{ij}+b_{ij}\right]$ with $\substack{i=1,2,\cdots,n \\ j=1,2,\cdots,m}$.

We designate the special matrix with all zero entries, called the **zero matrix**, by O. For example $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is the zero matrix in $M_{2\times3}(\mathbb{R})$. If $A = \begin{bmatrix} 3 & 4 & -1 \\ 2 & -7 & 0 \end{bmatrix}$ then

$$A+O = \begin{bmatrix} 3 & 4 & -1 \\ 2 & -7 & 0 \end{bmatrix}+\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 4 & -1 \\ 2 & -7 & 0 \end{bmatrix}$$

The dimension of O is usually clear from the context. We could however, if really needed, write $O_{2\times3}$ to indicate $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, or $O_{3\times5} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ etc.

Clearly $$A+O = A.$$

The **opposite** of a matrix A is designated by $-A$ ('minus A'). It is the matrix such that

$$A+(-A) = O$$

For example if $A = \begin{bmatrix} 3 & 4 & -1 \\ 2 & -7 & 0 \end{bmatrix}$ then $-A = \begin{bmatrix} -3 & -4 & 1 \\ -2 & 7 & 0 \end{bmatrix}$.

Each matrix has exactly one opposite, and it is clear that to get its opposite we only need to reverse the signs of all entries.

Note that matrix addition is in sync with our vector addition: $\begin{bmatrix} a \\ b \end{bmatrix}+\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} a+A \\ b+B \end{bmatrix}$ and $\begin{bmatrix} a \\ b \\ c \end{bmatrix}+\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} a+A \\ b+B \\ c+C \end{bmatrix}$

(2) Scalar multiplication

We can multiply any matrix from $M_{n\times m}(\mathbb{R})$ by a scalar. In this course we will only deal with real numbers as scalars. Scalar multiplication also goes component wise, e.g.

$$5\cdot\begin{bmatrix}3 & 4 & -1\\ 2 & -7 & 0\end{bmatrix} = \begin{bmatrix}15 & 20 & -5\\ 10 & -35 & 0\end{bmatrix}$$

In general we have

$$t\cdot\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1m}\\ a_{21} & a_{22} & \dots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nm}\end{bmatrix} = \begin{bmatrix}t\cdot a_{11} & t\cdot a_{12} & \dots & t\cdot a_{1m}\\ t\cdot a_{21} & t\cdot a_{22} & \dots & t\cdot a_{2m}\\ \vdots & \vdots & & \vdots\\ t\cdot a_{n1} & t\cdot a_{n2} & \dots & t\cdot a_{nm}\end{bmatrix}$$

Note that $0\cdot A = O$

$$0\cdot A = 0\cdot\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1m}\\ a_{21} & a_{22} & \dots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nm}\end{bmatrix} = \begin{bmatrix}0 & 0 & \dots & 0\\ 0 & 0 & \dots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \dots & 0\end{bmatrix} = O = O_{n\times m}$$

and $1\cdot A = A$

$$1\cdot A = 1\cdot\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1m}\\ a_{21} & a_{22} & \dots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nm}\end{bmatrix} = \begin{bmatrix}a_{11} & a_{12} & \dots & a_{1m}\\ a_{21} & a_{22} & \dots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nm}\end{bmatrix} = A$$

and finally $-A = -1\cdot A$

$$-1\cdot A = -1\cdot\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1m}\\ a_{21} & a_{22} & \dots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nm}\end{bmatrix} = \begin{bmatrix}-a_{11} & -a_{12} & \dots & -a_{1m}\\ -a_{21} & -a_{22} & \dots & -a_{2m}\\ \vdots & \vdots & & \vdots\\ -a_{n1} & -a_{n2} & \dots & -a_{nm}\end{bmatrix} = -A$$

Note that the scalar multiplication is in sync with the scalar multiplication of vectors: $t\cdot\begin{bmatrix}a\\ b\end{bmatrix} = \begin{bmatrix}t\cdot a\\ t\cdot b\end{bmatrix}$ and $t\cdot\begin{bmatrix}a\\ b\\ c\end{bmatrix} = \begin{bmatrix}t\cdot a\\ t\cdot b\\ t\cdot c\end{bmatrix}$

(3) Properties of matrix addition and scalar multiplication

The following properties are easily checked:

Theorem 11.1: If $A, B, C \in M_{n\times m}(\mathbb{R})$, and $s, t \in \mathbb{R}$ then

(a) $A + B = B + A$ [Commutativity]

(b) $A + (B + C) = (A + B) + C$ [Associativity]

(c) $A + O = A$

(d) $A + (-A) = O$

(e) $1 \cdot A = A$

(f) $s \cdot (t \cdot A) = (s \cdot t)\, A$

(g) $t \cdot (A + B) = t \cdot A + t \cdot B$ [Distributive law 1]

(h) $(t + s) \cdot A = t \cdot A + s \cdot A$ [Distributive law 2]

Furthermore

(i) $0 \cdot A = O$

(j) $-A = -1 \cdot A$

In a course on Linear Algebra you will see that properties (a) – (h) are *the* defining properties of a vector space. In fact $\mathbb{R}^2$, $\mathbb{R}^3$ and $M_{n\times m}(\mathbb{R})$ are very important examples of vector spaces.

(4) Matrix multiplication

We will define a matrix multiplication next. We begin with an example:
Let A be a 2×3 matrix and B a 3×4 matrix, e.g.

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

then

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

$$= \begin{bmatrix} 1 & 0 & 2 & 3 \\ 8 & -1 & 4 & 2 \end{bmatrix}$$

For those who have never seen matrix multiplication this needs some explanation.

The product $A\,B$ is computed as follows:

(1) To get the entry in row 1 column 1 we take row 1 of the first matrix and column 1 of the second matrix and compute what looks like a dot product:

$$\begin{bmatrix} \boxed{3 \quad 4 \quad -1} \\ 2 \quad 1 \quad 0 \end{bmatrix}\begin{bmatrix} \boxed{\begin{matrix}5\\-2\\6\end{matrix}} & \begin{matrix}-1 & 3 & 0\\ 1 & -2 & 2\\ 1 & -1 & 5\end{matrix}\end{bmatrix} = \begin{bmatrix} 3\cdot 5+4\cdot(-2)+(-1)\cdot 6 & 0 & 2 & 3 \\ 8 & -1 & 4 & 2 \end{bmatrix}$$

(2) To get the entry in row 1 column 2 we take row 1 of the first matrix and column 2 of the second matrix and compute what looks like a dot product:

$$\begin{bmatrix} \boxed{3 \quad 4 \quad -1} \\ 2 \quad 1 \quad 0 \end{bmatrix}\begin{bmatrix} \begin{matrix}5\\-2\\6\end{matrix} & \boxed{\begin{matrix}-1\\1\\1\end{matrix}} & \begin{matrix}3 & 0\\ -2 & 2\\ -1 & 5\end{matrix}\end{bmatrix} = \begin{bmatrix} 1 & 3\cdot(-1)+4\cdot 1+(-1) & 2 & 3 \\ 8 & -1 & 4 & 2 \end{bmatrix}$$

Etc.

(3) The last entry, row 2 column 4, is produced by taking row 2 of the first matrix and column 4 of the second matrix and computing what looks like a dot product:

$$\begin{bmatrix} 3 \quad 4 \quad -1 \\ \boxed{2 \quad 1 \quad 0} \end{bmatrix}\begin{bmatrix} \begin{matrix}5 & -1 & 3\\ -2 & 1 & -2\\ 6 & 1 & -1\end{matrix} & \boxed{\begin{matrix}0\\2\\5\end{matrix}}\end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 8 & -1 & 4 & 2\cdot 0+1\cdot 2+0\cdot 5 \end{bmatrix}$$

(4) Doing this for all entries we get:

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

$$\begin{bmatrix} 3\cdot 5+4\cdot(-2)+(-1)\cdot 6 & 3\cdot(-1)+4\cdot 1+(-1)\cdot 1 & 3\cdot 3+4\cdot(-2)+(-1)\cdot(-1) & 3\cdot 0+4\cdot 2+(-1)\cdot 5 \\ 2\cdot 5+1\cdot(-2)+0\cdot 6 & 2\cdot(-1)+1\cdot 1+0\cdot 1 & 2\cdot 3+1\cdot(-2)+0\cdot(-1) & 2\cdot 0+1\cdot 2+0\cdot 5 \end{bmatrix}$$

In general, when A is an $n\times m$ matrix and B an $m\times p$ matrix then

$$A\,B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mp} \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{bmatrix}$$

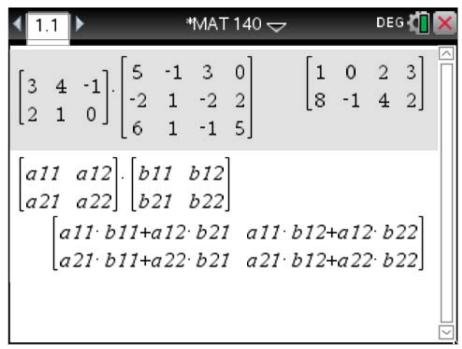
where $c_{ij} = a_{i1}\cdot b_{1j} + a_{i2}\cdot b_{2j} + a_{i3}\cdot b_{3j} + \cdots + a_{im}\cdot b_{mj} = \sum_{t=1}^{m} a_{it}\cdot b_{tj}$ or in compact form $\quad A\,B = \left[a_{ij}\right]\cdot\left[b_{ij}\right] = \left[\sum_{t=1}^{m} a_{it}\cdot b_{tj}\right].$

Note that to multiply two matrices we need the number of columns of the first matrix to be the same as the number of rows of the second:

$$A\,B = \begin{array}{c} \longleftarrow m \longrightarrow \\ \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \end{array} \begin{array}{c} \big\uparrow \\ m \\ \big\downarrow \end{array} \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mp} \end{bmatrix}$$

The $n\times m$ matrix A and $m\times p$ matrix B produce the $n\times p$ matrix $A\,B$.

The TI-Nspire computes a matrix product in exactly the way we described above:



Let $I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$ be the $n\times n$ matrix with 1s on the main diagonal and 0s off the main diagonal. Note that if A is an $n\times m$ matrix then $A\,I_m = A = I_n\,A$

$$A = \begin{array}{c} \longleftarrow m \longrightarrow \\ \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \end{array} \begin{array}{c} \big\uparrow \\ m \\ \big\downarrow \end{array} \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = \begin{array}{c} \longleftarrow n \longrightarrow \\ \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \end{array} \begin{array}{c} \big\uparrow \\ n \\ \big\downarrow \end{array} \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

One might wonder why matrix multiplication is defined the way it is. It would have been so much easier to do it component wise. The reason of the actual definition lies in the fact that we use matrices to perform linear transformations, e.g. $T(\vec{x}) = A\vec{x}$. If we have two linear transformations $T(\vec{x}) = A\vec{x}$ and $S(\vec{x}) = B\vec{x}$ it turns out that the composition of these two transformations $S\circ T$, can be expressed using a matrix multiplication precisely when we define it the way we did:

$$(T\circ S)(\vec{x}) = T(S(\vec{x})) = AB\,\vec{x}$$

But we will come back to this later.

(5) Properties of Matrix multiplication

Here are some useful properties:

Theorem 11.2: Let A, B and C be matrices (of the appropriate sizes, so that the indicated matrix operations are properly defined), then

(a) $A(BC) = (AB)C$ [Associativity]

(b) $A(B+C) = AB + AC$ [Distributative law 1]

(c) $(B+C)A = BA + CA$ [Distributative law 2]

(d) $(tA)B = A(tB) = t(AB)$

(e) $AI_m = A = I_n A$ **[** where A is an $n \times m$ matrix]

(f) $AO_{m\times p} = O_{n\times p}$ and $O_{p\times n} A = O_{p\times m}$

The proofs of **(b)**, **(c)**, **(d)**, **(e)** and **(f)** are pretty straight forward. The only non-trivial one is **(a)**, which is a bit messy to prove algebraically, but doable. It turns out to be much easier to prove once we have established the connection between matrix multiplications and composition of functions. This we will do in the near future.

Here are three important properties that we know to be true for real numbers, but that are false for matrices:

(a) If $x, y \in \mathbb{R}$ then $x \cdot y = y \cdot x$
(b) If $x, y \in \mathbb{R}$ then $x \cdot y = 0$ implies $x = 0$ or $y = 0$
(c) If $a, x, y \in \mathbb{R}$ and $a \neq 0$ then $ax = ay$ implies $x = y$

For matrices these properties are **not** true in general:

(a) If A, B are matrices then in general: $\boxed{AB \neq BA}$

Here are three scenarios of what can go wrong:

(i) If A is an $n \times m$ matrix and B is an $m \times p$ then AB is an $n \times p$ matrix, but BA does not even exist when $n \neq p$.

But even if they are the right size ($n = p$) …

(ii) If A is an $n \times m$ matrix and B is an $m \times n$ then AB is an $n \times n$ matrix, but although BA exists in this case it is an $m \times m$ matrix, i.e. $AB \neq BA$ if $m \neq n$.

But even if they are both square matrices ($m = n$)…

(iii) If both A and B are $n\times n$ matrices we still usually have $AB \neq BA$:

e.g. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}$ but $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ -1 & -2 \end{bmatrix}$

Note that this does **not** mean that matrices never commute. It does happen sometimes that

$AB = BA$: e.g. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

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

(b) If A, B are matrices then $AB = O$ does **not** imply $A = O$ or $B = O$

For example: $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

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

(c) If A, B, C, with $B \neq O$, are matrices then $AB = CB$ does **not** imply $A = C$.

For example: $\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 6 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}}_{B} = \begin{bmatrix} 3 & 6 \\ 8 & 16 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix}}_{C}\underbrace{\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}}_{B}$ but $\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 6 \end{bmatrix}}_{A} \neq \underbrace{\begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix}}_{C}$

(6) Powers of matrices

We call an $n\times n$ matrix a **square** matrix. If A is a square matrix we can define

$A^2 = AA,\ A^3 = AAA,\ A^4 = AAAA$ etc. In general defined recursively:

Definition: **(a)** $A^0 = I_n,\ A^1 = A$

(b) $A^n = A^{n-1}A$ when $n = 1, 2, 3\cdots$

The following rules then hold for powers of matrices:

Theorem 11.3: **(a)** $A^n A^m = A^{n+m}$ for $n,\ m \in \mathbb{Z}_0^+$ [in fact $n,\ m \in \mathbb{Z}$]

(b) $(A^n)^m = A^{nm}$ for $n,\ m \in \mathbb{Z}_0^+$ [in fact $n,\ m \in \mathbb{Z}$]

For example: $*$ $A^3\cdot A^4 = (AAA)(AAAA) = A^7 = A^{3+4}$

$*$ $(A^2)^3 = (AA)(AA)(AA) = A^6 = A^{2\cdot 3}$

(7) The transpose

The **transpose** of an $n \times m$ matrix A is the $m \times n$ matrix denoted by A^{T} which is formed from A by transposing rows and columns: i.e. the i^{th} column becomes the i^{th} row, and consequently the j^{th} row becomes the j^{th} column

$$\text{If} \quad A = \begin{bmatrix} \boxed{\begin{matrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{matrix}} & \boxed{\begin{matrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \\ \cdots \end{matrix} & \boxed{\begin{matrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{matrix}} \end{bmatrix} \quad \text{then} \quad A^{\mathsf{T}} = \begin{bmatrix} \boxed{a_{11} \quad a_{21} \quad \cdots \quad a_{n1}} \\ \boxed{a_{12} \quad a_{22} \quad \cdots \quad a_{n2}} \\ \\ \boxed{a_{1m} \quad a_{2m} \quad \cdots \quad a_{nm}} \end{bmatrix}$$

$$\text{For example:} \quad A = \begin{bmatrix} 3 & 4 & -1 \\ 2 & -7 & 0 \end{bmatrix} \quad \text{then} \quad A^{\mathsf{T}} = \begin{bmatrix} 3 & 2 \\ 4 & -7 \\ -1 & 0 \end{bmatrix}$$

Here are some important properties:

Theorem 11.4: **(a)** $(A^{\mathsf{T}})^{\mathsf{T}} = A$

(b) $(A+B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}}$

(c) $(tA)^{\mathsf{T}} = t(A^{\mathsf{T}})$

(d) $(tA+sB)^{\mathsf{T}} = tA^{\mathsf{T}} + sB^{\mathsf{T}}$ [combining **(b)** and **(c)**]

(e) $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$

The proofs of **(a) – (d)** are pretty straight forward. The only non trivial case is **(e)** the proof of which is algebraically messy, but straightforward.

(8) The inverse of a 2×2 matrix

We will discuss inverses of matrices in general more extensively in another section. Here we'll give the formal definition, the uniqueness theorem, and examine the 2×2 case.

An $n \times n$ matrix A is called **invertible** if there exist an $n \times n$ matrix B such that

$$AB = I_n = BA$$

$$\text{For example:} \quad \begin{bmatrix} 3 & 1 & 3 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & -5 \\ 1 & 0 & -3 \\ -2 & 1 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -5 \\ 1 & 0 & -3 \\ -2 & 1 & 6 \end{bmatrix} \begin{bmatrix} 3 & 1 & 3 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Theorem 11.5: If an $n \times n$ matrix is invertible its inverse is unique.

Proof: Suppose A has two inverses B and C: i.e.

$AB = I_n = BA$ and $AC = I_n = CA$, then

$B = BI_n = B(AC) = (BA)C = I_n C = C$ hence $B = C$. □

Since inverses are unique we can use a unique notation for *the* inverse of a matrix A namely: A^{-1}. Hence

$$AA^{-1} = I_n = A^{-1}A$$

The inverse of a 2×2 matrix is particularly easy to compute

Theorem 11.6: A 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$ in which case its inverse is given by:

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Proof: Clearly if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ with $ad - bc \neq 0$ then

$$BA = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$AB = \frac{1}{ad-bc}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So that $B = A^{-1}$ and hence $A^{-1} = \dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

The above proves that $ad - bc \neq 0$ is a *sufficient* condition for invertibility. To prove $ad - bc \neq 0$ is also a necessary condition, let's assume that $ad - bc = 0$ [i.e. $ad = bc$], and try to find an inverse B for A. Let $B = \begin{bmatrix} r & s \\ t & u \end{bmatrix}$. We would need: $AB = I_n = BA$ i.e.

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} r & s \\ t & u \end{bmatrix} = \begin{bmatrix} ar+bt & as+bu \\ cr+dt & cs+du \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

i.e. $\begin{cases} ar+bt=1 \\ as+bu=0 \\ cr+dt=0 \\ cs+du=1 \end{cases} \Rightarrow \begin{cases} adr+bdt=d \\ ads+bdu=0 \\ bcr+bdt=0 \\ bcs+bdu=b \end{cases}$ $\begin{matrix} d=0 \\ b=0 \end{matrix}$ using $ad=bc$, but this is

impossible since then $BA = \begin{bmatrix} r & s \\ t & u \end{bmatrix}\begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} = \begin{bmatrix} * & 0 \\ * & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. $\square$

Example: $A = \begin{bmatrix} 4 & 2 \\ 3 & 2 \end{bmatrix} \Rightarrow A^{-1} = \dfrac{1}{4\cdot 2 - 3\cdot 2}\begin{bmatrix} 2 & -2 \\ -3 & 4 \end{bmatrix} = \dfrac{1}{2}\begin{bmatrix} 2 & -2 \\ -3 & 4 \end{bmatrix}$ and indeed

$\begin{bmatrix} 4 & 2 \\ 3 & 2 \end{bmatrix}\left(\dfrac{1}{2}\begin{bmatrix} 2 & -2 \\ -3 & 4 \end{bmatrix}\right) = \dfrac{1}{2}\begin{bmatrix} 4 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & -2 \\ -3 & 4 \end{bmatrix} = \dfrac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and

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

(9) Determinant of a 2×2 matrix.

The **determinant** of a 2×2 matrix is defined as follows:

Definition: $\det(A) = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} := ad - bc$.

Theorem 11.7: A matrix A is invertible if and only if $\det(A) \neq 0$.

The proof for this will be postponed till later, but the previous theorem at least proves the theorem for the 2×2 case, as well as giving us the nice formula

$$A^{-1} = \frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

12. Determinants

Determinants can be introduced in a number of ways. Since we are focusing on $\mathbb{R}^2$ and $\mathbb{R}^3$ in this course, we are mainly interested in 2×2, 3×3 (and 4×4) determinants, and hence we do not need to worry about defining determinants in general. In the next course on Linear Algebra we'll define it in general for any $n\times n$ matrix. But here we'll start by defining the determinants of 1×1, 2×2 and 3×3 matrices. Of course from these definitions you can see how we could proceed in general.

Definition: The determinants of 1×1, 2×2 and 3×3 matrices are defined as follows:

a) $\det\begin{bmatrix} u_1 \end{bmatrix} = u_1$

b) $\det\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} = u_1 v_2 - u_2 v_1 = u_1 \det[v_2] - u_2 \det[v_1]$

c) $\det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = u_1 \det\begin{bmatrix} v_2 & w_2 \\ v_3 & w_3 \end{bmatrix} - u_2 \det\begin{bmatrix} v_1 & w_1 \\ v_3 & w_3 \end{bmatrix} + u_3 \det\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix}$

Example 1:

a) $\det\begin{bmatrix} 7 \end{bmatrix} = 7$

b) $\det\begin{bmatrix} 6 & 1 \\ 9 & 5 \end{bmatrix} = 6\cdot 5 - 9\cdot 1 = 21$

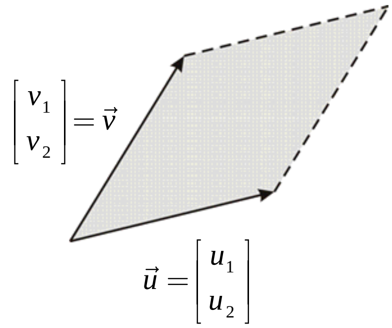
c) $\det\begin{bmatrix} 3 & 4 & 8 \\ 7 & 6 & 1 \\ 2 & 9 & 5 \end{bmatrix} = 3\det\begin{bmatrix} 6 & 1 \\ 9 & 5 \end{bmatrix} - 7\det\begin{bmatrix} 4 & 8 \\ 9 & 5 \end{bmatrix} + 2\det\begin{bmatrix} 4 & 8 \\ 6 & 1 \end{bmatrix} = 339$

We will now derive a host of very interesting properties. In the next Linear Algebra course we will see that most of these properties hold in general, but here we will prove them only for 1×1, 2×2 and 3×3 matrices.

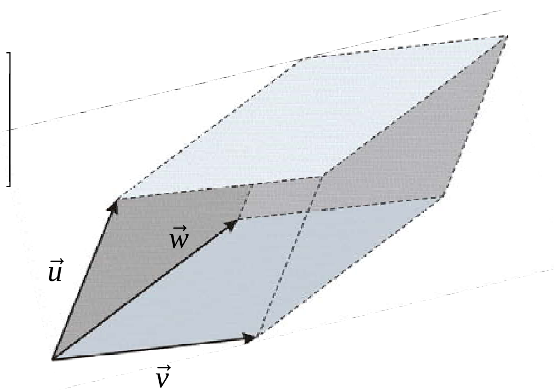
First of all, determinants appear to be related to areas (in $\mathbb{R}^2$) and volumes (in $\mathbb{R}^3$):

Theorem 12.1: (Areas and volumes)

(a) The area of the parallelogram spanned by the vector $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is given by $\left| \det \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \right|$.

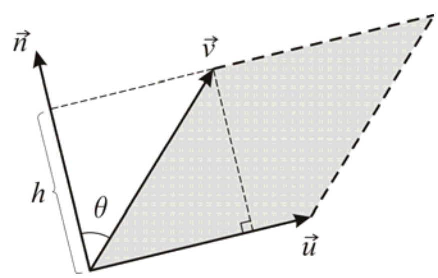


(b) The volume of the parallelepiped spanned by the vectors $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ is given by $\left| \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \right|$.



Proof:

(a) Note that $\vec{n} = \begin{bmatrix} -u_2 \\ u_1 \end{bmatrix}$ is perpendicular to $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, *and* $\|\vec{u}\| = \|\vec{n}\|$.



Hence $h = \left\| \mathrm{proj}_{\vec{n}}(\vec{v}) \right\| = \left\| \frac{\vec{n}\cdot\vec{v}}{\vec{n}\cdot\vec{n}}\vec{n} \right\| = \frac{|u_1v_2 - u_2v_1|}{\|\vec{n}\|} = \frac{|u_1v_2 - u_2v_1|}{\|\vec{u}\|}$

[or alternatively $h = \|\vec{v}\| \cdot |\cos\theta| = \frac{\|\vec{n}\| \cdot \|\vec{v}\| \cdot |\cos\theta|}{\|\vec{n}\|} = \frac{|\vec{n}\cdot\vec{v}|}{\|\vec{u}\|}$]

and thus the area of the parallelogram is computed as follows

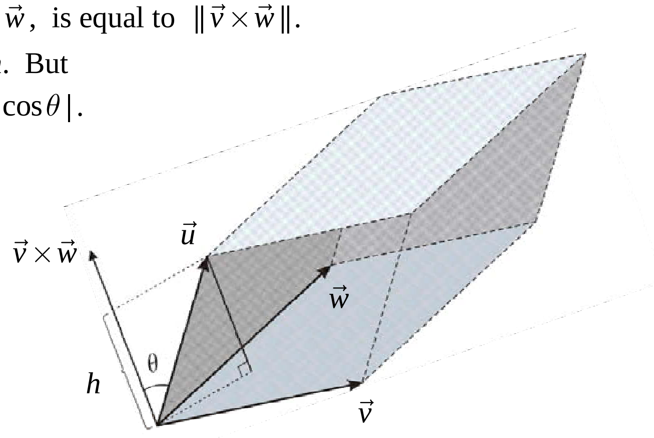
$$\mathcal{A} = \textit{height} \cdot \textit{base} = h \cdot \|\vec{u}\| = |u_1v_2 - u_2v_1| = \left| \det \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \right|$$

(b) Note that by Theorem 7.3 the area of the base of the parallelepiped, i.e. the area of the parallelogram spanned by $\vec{v}$ and $\vec{w}$, is equal to $\|\vec{v}\times\vec{w}\|$.

We just need to compute the height h. But trigonometry tells us that $h = \|\vec{u}\|\cdot|\cos\theta|$.

Hence since the volume can be computed by:

$$\mathcal{V} = height\cdot base$$



We find:

$$\mathcal{V} = height\cdot base = \big|\,\|\vec{u}\|\cdot\|\vec{v}\times\vec{w}\|\cos\theta\,\big| = \big|\vec{u}\cdot(\vec{v}\times\vec{w})\big|$$

$$= \left|\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}\cdot\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\times\begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix}\right)\right| = \left|\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}\cdot\left(\begin{bmatrix}v_2w_3-v_3w_2\\v_3w_1-v_1w_3\\v_1w_2-v_2w_1\end{bmatrix}\right)\right|$$

$$= |u_1(v_2w_3-v_3w_2)+u_2(v_3w_1-v_1w_3)+u_3(v_1w_2-v_2w_1)|$$

$$= \left|u_1\det\begin{bmatrix}v_2 & w_2\\v_3 & w_3\end{bmatrix} - u_2\det\begin{bmatrix}v_1 & w_1\\v_3 & w_3\end{bmatrix} + u_3\det\begin{bmatrix}v_1 & w_1\\v_2 & w_2\end{bmatrix}\right|$$

$$= \left|\det\begin{bmatrix}u_1 & v_1 & w_1\\u_2 & v_2 & w_2\\u_3 & v_3 & w_3\end{bmatrix}\right| \qquad \square$$

Note 1: The expression $\vec{u}\cdot(\vec{v}\times\vec{w})$ is sometimes referred to as the **triple product**, and as the above proof shows: $\vec{u}\cdot(\vec{v}\times\vec{w}) = \det\begin{bmatrix}u_1 & v_1 & w_1\\u_2 & v_2 & w_2\\u_3 & v_3 & w_3\end{bmatrix}$

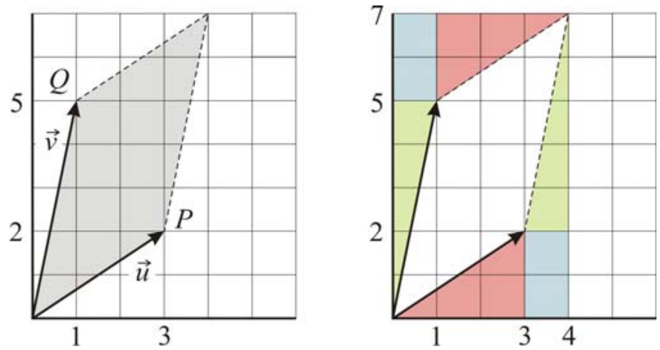
Note 2: We can derive Theorem 12.1 (a) also in another way: The areas of the two parallelograms, one spanned by $\begin{bmatrix}u_1\\u_2\end{bmatrix}$ & $\begin{bmatrix}v_1\\v_2\end{bmatrix}$ *and* the other by $\begin{bmatrix}u_1\\u_2\\0\end{bmatrix}$ & $\begin{bmatrix}v_1\\v_2\\0\end{bmatrix}$, are equal.

Hence by Theorem 7.3, this area is equal to

$$\left\|\begin{bmatrix}u_1\\u_2\\0\end{bmatrix}\times\begin{bmatrix}v_1\\v_2\\0\end{bmatrix}\right\| = \left\|\begin{bmatrix}0\\0\\u_1v_2-u_2v_1\end{bmatrix}\right\| = |u_1v_2-u_2v_1|$$

Example 2: The area of the parallelogram spanned by $\vec{u} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$ is

$$\det \begin{bmatrix} 3 & 1 \\ 2 & 5 \end{bmatrix} = 15 - 2 = 13 .$$



White area is
$28 - 6 - 5 - 4 = 13$

Example 3: The volume of the parallelepiped spanned by $\begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is given by

$$\left| \det \begin{bmatrix} 3 & 1 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = 13 .$$

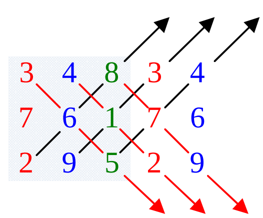
[Draw a picture to see how this is related to example 2]

Note 3: Sometimes the following *cute* way of computing a 3×3 determinant is shown,

$$-2 \cdot 6 \cdot 8 - 9 \cdot 1 \cdot 3 - 5 \cdot 7 \cdot 4 = -263$$

$$\det \begin{bmatrix} 3 & 4 & 8 \\ 7 & 6 & 1 \\ 2 & 9 & 5 \end{bmatrix} = 339$$

and



$\det(A) = 602 - 263 = 339$

$$3 \cdot 6 \cdot 5 + 4 \cdot 1 \cdot 2 + 8 \cdot 7 \cdot 9 = 602$$

This follows from our definition:

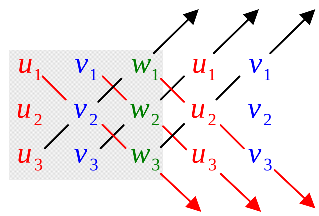
$$\det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = u_1 \det \begin{bmatrix} v_2 & w_2 \\ v_3 & w_3 \end{bmatrix} - u_2 \det \begin{bmatrix} v_1 & w_1 \\ v_3 & w_3 \end{bmatrix} + u_3 \det \begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix}$$

$$= u_1 v_2 w_3 - u_1 v_3 w_2 - u_2 v_1 w_3 + u_2 v_3 w_1 + u_3 v_1 w_2 - u_3 v_2 w_1$$

$$= (u_1 v_2 w_3 + u_2 v_3 w_1 + u_3 v_1 w_2) + (-u_1 v_3 w_2 - u_2 v_1 w_3 - u_3 v_2 w_1)$$

Just compare this with:

$$-u_3v_2w_1 - v_3w_2u_1 - w_3u_2v_1$$



$$u_1v_2w_3 + v_1w_2u_3 + w_1u_2v_3$$

Unfortunately this doesn't generalize in an obvious way to bigger matrices. So it is cute but not very useful.

Cofactor expansion

Let's examine the definition of 3×3 determinants a little: Observe how we get the determinant by '**expanding**' along the first column:

$$\det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = u_1 \det\begin{bmatrix} v_2 & w_2 \\ v_3 & w_3 \end{bmatrix} - u_2 \det\begin{bmatrix} v_1 & w_1 \\ v_3 & w_3 \end{bmatrix} + u_3 \det\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix}$$

Let's explain what we mean by 'expanding' along the first column:

To get the first term $u_1 \det\begin{bmatrix} v_2 & w_2 \\ v_3 & w_3 \end{bmatrix}$ we take the first element in column 1 u_1 cross out the row and column it is in and take the determinant of the remaining elements

$$\begin{bmatrix} u_1 & & \\ & v_2 & w_2 \\ & v_3 & w_3 \end{bmatrix} \quad \text{to get the product} \quad u_1 \det\begin{bmatrix} v_2 & w_2 \\ v_3 & w_3 \end{bmatrix}$$

Next, take the second element in column 1 u_2 cross out the row and column it is in and take the determinant of the remaining elements, *and* throw in an extra minus sign

$$\begin{bmatrix} & v_1 & w_1 \\ u_2 & & \\ & v_3 & w_3 \end{bmatrix} \quad \text{to get the product} \quad -u_2 \det\begin{bmatrix} v_1 & w_1 \\ v_3 & w_3 \end{bmatrix}$$

Finally, take the third element in column 1 u_3 cross out the row and column it is in and take the determinant of the remaining elements

$$\begin{bmatrix} & v_1 & w_1 \\ & v_2 & w_2 \\ u_3 & & \end{bmatrix} \quad \text{to get the product} \quad u_3 \det\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix}$$

Let's put the pictures together, to 'see' the *expansion* along column 1:

$$\det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = u_1 \det\begin{bmatrix} v_2 & w_2 \\ v_3 & w_3 \end{bmatrix} - u_2 \det\begin{bmatrix} v_1 & w_1 \\ v_3 & w_3 \end{bmatrix} + u_3 \det\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix}$$

$$+\begin{bmatrix} u_1 & & \\ & v_2 & w_2 \\ & v_3 & w_3 \end{bmatrix} \quad - \begin{bmatrix} & v_1 & w_1 \\ u_2 & & \\ & v_3 & w_3 \end{bmatrix} \quad + \begin{bmatrix} & v_1 & w_1 \\ & v_2 & w_2 \\ u_3 & & \end{bmatrix}$$

The interesting thing is that we can expand along any row or column and get the same value, as long as we put the $+$ and $-$ signs according to the following table:

$$\begin{array}{|ccccc} \hline + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{array}$$

For example, let's see what expansion along the second column would look like:

$$\det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = -v_1 \det\begin{bmatrix} u_2 & w_2 \\ u_3 & w_3 \end{bmatrix} w_3 + v_2 \det\begin{bmatrix} u_1 & w_1 \\ u_3 & w_3 \end{bmatrix} - v_3 \det\begin{bmatrix} u_1 & w_1 \\ u_2 & w_2 \end{bmatrix}$$

That this is correct follows from the definition of the determinant and some algebra:

$$\det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = u_1v_2w_3 - u_1v_3w_2 - u_2v_1w_3 + u_2v_3w_1 + u_3v_1w_2 - u_3v_2w_1 \qquad \text{[by definition]}$$

$$= -v_1(u_2w_3 - u_3w_2) + v_2(u_1w_3 - u_3w_1) - v_3(u_1w_2 - u_2w_1)$$

$$= -v_1\det\begin{bmatrix} u_2 & w_2 \\ u_3 & w_3 \end{bmatrix} w_3 + v_2\det\begin{bmatrix} u_1 & w_1 \\ u_3 & w_3 \end{bmatrix} - v_3\det\begin{bmatrix} u_1 & w_1 \\ u_2 & w_2 \end{bmatrix}$$

Hence indeed using the $+$ and $-$ signs as indicated by the alternating signs diagram we can expand along the second column as follows:

$$\det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = -v_1\det\begin{bmatrix} u_2 & w_2 \\ u_3 & w_3 \end{bmatrix} + v_2\det\begin{bmatrix} u_1 & w_1 \\ u_3 & w_3 \end{bmatrix} - v_3\det\begin{bmatrix} u_1 & w_1 \\ u_2 & w_2 \end{bmatrix}$$

$$-\begin{bmatrix} — & v_1 & — \\ u_2 & | & w_2 \\ u_3 & | & w_3 \end{bmatrix} \qquad + \begin{bmatrix} u_1 & | & w_1 \\ — & v_2 & — \\ u_3 & | & w_3 \end{bmatrix} \qquad - \begin{bmatrix} u_1 & | & w_1 \\ u_2 & | & w_2 \\ — & v_3 & — \end{bmatrix}$$

where the $-$ $+$ $-$ come from the places v_1, v_2 and v_3 are on in the $+$ and $-$ table:

$$\begin{array}{|ccccc} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{array}$$

It is easy to check that this also works for rows too: e.g. expansion along row 3

$$u_3\det\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix} - v_3\det\begin{bmatrix} u_1 & w_1 \\ u_2 & w_2 \end{bmatrix} + w_3\det\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} =$$

$$= u_3(v_1w_2 - v_2w_1) - v_3(u_1w_2 - u_2w_1) + w_3(u_1v_2 - u_2v_1)$$

$$= u_3v_1w_2 - u_3v_2w_1 - u_1v_3w_2 + u_2v_3w_1 + u_1v_2w_3 - u_2v_1w_3 = \det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$$

The determinants of the above *sub-matrices* are called **minors**: e.g. the minor of a_{23} in

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \text{is} \quad M_{23} = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix}$$

We call $(-1)^{i+j}M_{ij}$ the **cofactor of** a_{ij} i.e. the minor with the $\pm$ sign: e.g.

$$(-1)^{2+3}M_{23} = -\det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix}$$

Theorem 12.2: **Cofactor expansion**

The $\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ can be calculated by expanding along *any* row or *any* column:

[These expansions are called *cofactor expansions* along a row or column]

Along the i th row:

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = (-1)^{i+1}a_{i1}M_{i1} + (-1)^{i+2}a_{i2}M_{i2} + (-1)^{i+3}a_{i3}M_{i3}$$

Along the j th column:

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = (-1)^{1+j}a_{1j}M_{1j} + (-1)^{2+j}a_{2j}M_{2j} + (-1)^{3+j}a_{3j}M_{3j}$$

Proof: Our discussion preceding this theorem showed that the cofactor expansions along column 1, column 2 and row 3 are all equal. The cofactor expansions along the other rows and column can be shown to be equal to the other three in a similar way [It is easy to check all 6 cases (3 columns and 3 rows) by hand]; or we could invoke the next theorem about the determinants of matrices and their transposes. This theorem is true in general for $n\times n$ matrix, but it takes more work to prove this, which will be done in the next Linear Algebra course. $\square$

Example 4:

$$\det\begin{bmatrix} 3 & 4 & 8 \\ 7 & 6 & 1 \\ 2 & 9 & 5 \end{bmatrix} = (-1)^{2+1}a_{21}M_{21} + (-1)^{2+2}a_{22}M_{22} + (-1)^{2+3}a_{23}M_{23}$$

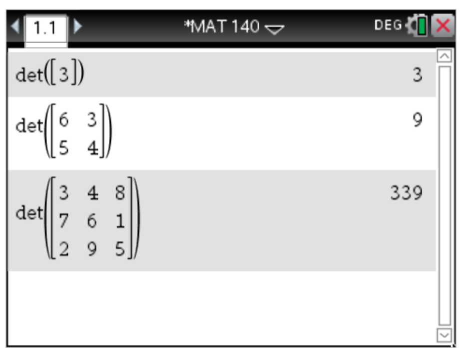
$$= -7\det\begin{bmatrix} 4 & 8 \\ 9 & 5 \end{bmatrix} + 6\det\begin{bmatrix} 3 & 8 \\ 2 & 5 \end{bmatrix} - 1\det\begin{bmatrix} 3 & 4 \\ 2 & 9 \end{bmatrix} = 339$$

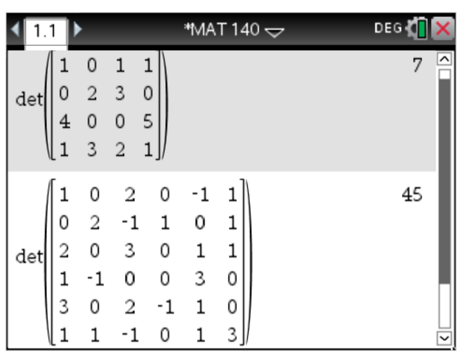
$$\det\begin{bmatrix} 3 & 4 & 8 \\ 7 & 6 & 1 \\ 2 & 9 & 5 \end{bmatrix} = (-1)^{1+3}a_{13}M_{13} + (-1)^{2+3}a_{23}M_{23} + (-1)^{3+3}a_{33}M_{33}$$

$$= 8\det\begin{bmatrix} 7 & 6 \\ 2 & 9 \end{bmatrix} - 1\det\begin{bmatrix} 3 & 4 \\ 2 & 9 \end{bmatrix} + 5\det\begin{bmatrix} 3 & 4 \\ 7 & 6 \end{bmatrix} = 339$$

The TI-Nspire

The TI-Nspire of course knows how to compute determinants, and not just of 1×1, 2×2 and 3×3 matrices, but of square matrices of any size. It's function is of course called: **det( )**





13. Some important Properties of Determinants

There is much to say about determinants. There is a great number of properties of determinants we could discuss, but we will just mention a few of the most important and useful properties--without any proofs. In the next Linear Algebra course we will cover many more properties and prove most of them in general.

Theorem 13.1: $\det(I_n) = 1$

Theorem 13.2: If A has a row (or column) of zeros then $\det(A) = 0$

Theorem 13.3: If A has two identical rows (or columns) then $\det(A) = 0$

Theorem 13.4: Adding a multiple of one row to another row, or adding a multiple of one column to another column doesn't change the determinant.

Theorem 13.5: Multiplying a row (or column) by a constant k changes the determinant by a factor of k.

Theorem 13.6: Interchanging two rows (or columns) changes the sign of the determinant.

Theorem 13.7: $\det(A^{T}) = \det(A)$

Theorem 13.8: $\det(AB) = \det(A)\det(B)$

Theorem 13.9: A is invertible if and only if $\det(A) \neq 0$

Theorem 13.10: If A is invertible then $\det(A^{-1}) = \dfrac{1}{\det(A)}$

Examples

1. $$\det\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} = 1\det\begin{bmatrix}1&0\\0&1\end{bmatrix} - 0\det\begin{bmatrix}0&0\\0&1\end{bmatrix} + 0\det\begin{bmatrix}0&0\\1&0\end{bmatrix} = 1$$

(Expanding along the first column)

2. $$\det\begin{bmatrix}3&4&5\\0&0&0\\1&7&2\end{bmatrix} = -0\det\begin{bmatrix}4&5\\7&2\end{bmatrix}+0\det\begin{bmatrix}3&5\\1&2\end{bmatrix}-0\det\begin{bmatrix}3&4\\1&7\end{bmatrix} = 0$$

(Expanding along the second row)

3. $$\det\begin{bmatrix}3&4&5\\2&7&8\\3&4&5\end{bmatrix} = -2\det\begin{bmatrix}4&5\\4&5\end{bmatrix}+7\det\begin{bmatrix}3&5\\3&5\end{bmatrix}-8\det\begin{bmatrix}3&4\\3&4\end{bmatrix} = 0$$

(Expanding along the second row)

4. $$\det\begin{bmatrix}2&1&3\\7&8&9\\4&5&6\end{bmatrix} = \det\begin{bmatrix}2&1&3\\7&8&9\\0&3&0\end{bmatrix} = -3\det\begin{bmatrix}2&3\\7&9\end{bmatrix} = 9$$

-2 Row 1 + Row 3

5. $$\det\begin{bmatrix}3&4&5\\5\cdot8&5\cdot6&5\cdot9\\1&7&2\end{bmatrix} = -5\cdot8\det\begin{bmatrix}4&5\\7&2\end{bmatrix}+5\cdot6\det\begin{bmatrix}3&5\\1&2\end{bmatrix}-5\cdot9\det\begin{bmatrix}3&4\\1&7\end{bmatrix}$$

$$= 5\cdot\left(-8\det\begin{bmatrix}4&5\\7&2\end{bmatrix}+6\det\begin{bmatrix}3&5\\1&2\end{bmatrix}-9\det\begin{bmatrix}3&4\\1&7\end{bmatrix}\right)$$

$$= 5\cdot\det\begin{bmatrix}3&4&5\\8&6&9\\1&7&2\end{bmatrix}$$

6. $$\det\begin{bmatrix}2&1&3\\7&8&9\\4&5&6\end{bmatrix} = -7\det\begin{bmatrix}1&3\\5&6\end{bmatrix}+8\det\begin{bmatrix}2&3\\4&6\end{bmatrix}-9\det\begin{bmatrix}2&1\\4&5\end{bmatrix}$$

$$\text{and} \quad \det\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = 7\det\begin{bmatrix} 1 & 3 \\ 5 & 6 \end{bmatrix} - 8\det\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} + 9\det\begin{bmatrix} 2 & 1 \\ 4 & 5 \end{bmatrix}$$

$$\text{hence indeed:} \quad \det\begin{bmatrix} 2 & 1 & 3 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{bmatrix} = -\det\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

7. $$\det\begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} = u_1\det\begin{bmatrix} v_2 & v_3 \\ w_2 & w_3 \end{bmatrix} - v_1\det\begin{bmatrix} u_2 & u_3 \\ w_2 & w_3 \end{bmatrix} + w_1\det\begin{bmatrix} u_2 & u_3 \\ v_2 & v_3 \end{bmatrix}$$

$$= u_1v_2w_3 - u_1w_2v_3 - v_1u_2w_3 + v_1w_2u_3 + w_1u_2v_3 - w_1v_2u_3$$

$$\det\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = u_1\det\begin{bmatrix} v_2 & w_2 \\ v_3 & w_3 \end{bmatrix} - u_2\det\begin{bmatrix} v_1 & w_1 \\ v_3 & w_3 \end{bmatrix} + u_3\det\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix}$$

$$= u_1v_2w_3 - u_1v_3w_2 - u_2v_1w_3 + u_2v_3w_1 + u_3v_1w_2 - u_3v_2w_1$$

8. $$\det\left(\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 2 & 3 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 16 & 21 \\ 30 & 39 \end{bmatrix}\right) = 16\cdot 39 - 30\cdot 21 = -6$$

$$\det\left(\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}\right)\det\left(\begin{bmatrix} 5 & 6 \\ 2 & 3 \end{bmatrix}\right) = (2\cdot 5 - 4\cdot 3)(5\cdot 3 - 2\cdot 6) = (-2)\cdot 3 = -6$$

9. $A^{-1} = \dfrac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. If $\det(A) = 0$ then A^{-1} doesn't exist.

10. When $A = \begin{bmatrix} 3 & 5 \\ 1 & 4 \end{bmatrix}$ then $A^{-1} = \dfrac{1}{7}\begin{bmatrix} 4 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 4/7 & -5/7 \\ -1/7 & 3/7 \end{bmatrix}$.

So that $\det\begin{bmatrix} 3 & 5 \\ 1 & 4 \end{bmatrix} = 7$ and

$$\det(A^{-1}) = \det\begin{bmatrix} 4/7 & -5/7 \\ -1/7 & 3/7 \end{bmatrix} = \frac{12}{49} - \frac{5}{49} = \frac{7}{49} = \frac{1}{7} = \frac{1}{\det(A)}$$

14. Reduced Row Echelon Form

The following matrix is in what we call reduced row echelon form:

$$\begin{bmatrix} 0 & 1 & 3 & \boxed{0} & 2 & 4 & \boxed{0} & 7 & 9 \\ 0 & 0 & 0 & 1 & 6 & 5 & \boxed{0} & 8 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

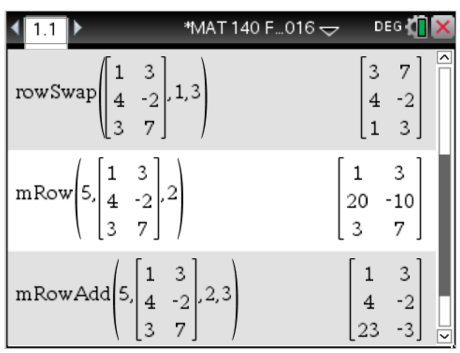
Definition 1: We say a matrix in **reduced row echelon form** if

(1) All rows with only zeros in it, if any, are at the bottom of the matrix.
(2) Reading from left to right, the first non-zero element in any non-zero row is a 1.
(These non-zero elements are called pivots, whether or not they are 1)
(3) If row i and row $i+1$ are two consecutive non-zero rows, then the pivot in row $i+1$ is to the right of the pivot in row i.
(4) Each column containing a pivot contains besides the pivot only zeros.

Any matrix can be transformed into reduced row echelon form, as we will see, by using just three types of operations, called **elementary row operations**:

Type I: Swapping two rows.
Type II: Multiplying a row with a non-zero constant.
Type III: Adding a multiple of one row to another.

These three elementary row operations can be performed by the TI-Nspire using the functions **rowSwap, mRow, and mRowAdd**.



We have to know how to do these operations by hand, but it is easy to make mistakes. Using the TI-Nspire we do not have to make any arithmetical errors. In fact it allows us to focus on what is going on—*how* the process we are going to look at works—and not on arithmetic and trying to avoid and tracing mistakes. Arithmetic is not at all the topic of this course. We assume you have mastered that. So to avoid mistakes … use the TI-Nspire!

Let's illustrate the procedure with the matrix $\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix}$.

(1) Swap row 1 and row 2 to get $\begin{bmatrix} 1 & 1 & 1 & 6 \\ 3 & 2 & -1 & 1 \\ 2 & 0 & 1 & -1 \end{bmatrix}$.

We have produced our first pivot 1. [We could have also multiplied row 1 by 1/3.]

(2) Add -3 times row 1 to row 2 to get $\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 2 & 0 & 1 & -1 \end{bmatrix}$

(3) Add -2 times row 1 to row 3 to get $\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 0 & -2 & -1 & -13 \end{bmatrix}$

(4) Multiply both row 2 and row 3 by -1 $\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 2 & 1 & 13 \end{bmatrix}$

(5) Add -2 times row 2 to row 3 to get $\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & -7 & -21 \end{bmatrix}$

(6) Multiply row 3 by $-\frac{1}{7}$ to get $\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & 1 & 3 \end{bmatrix}$

We have accomplished a lot. (**1**), (**2**) and (**3**) of the reduced row echelon properties are now satisfied. We just need to produce zeros above the pivots:

(7) Add -4 times row 3 to row 2 to get $\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$

(8) Add -1 times row 3 to row 1 to get $\begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$ and finally

(9) Add -1 times row 2 to row 1 to get $\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$

We have produced a reduced row echelon form.

It is useful to have a notation to report these elementary row operations:

Type I: swapping rows e.g. $\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 3 & 2 & -1 & 1 \\ 2 & 0 & 1 & -1 \end{bmatrix}.$

[This operation is done by the TI-Nspire as follows: **rowSwap(m, 1, 2)** where m is the first matrix]

Type II: Multiplying a row with a non zero constant , e.g.

$$\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & -7 & -21 \end{bmatrix} \xrightarrow{-\frac{1}{7}R_3} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

[This operation is done by the TI-Nspire as follows: **mRow(-1/7, m, 3)** where m is the first matrix]

Type III: Adding a multiple of one row to another, e.g.

$$\begin{bmatrix} 1 & 1 & 1 & 6 \\ 3 & 2 & -1 & 1 \\ 2 & 0 & 1 & -1 \end{bmatrix} \xrightarrow{-3R_1+R_2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 2 & 0 & 1 & -1 \end{bmatrix}.$$

[This operation is done by the TI-Nspire as follows: **mRowAdd(-3, m, 1, 2)** where m is the first matrix]

With this notation we can describe the steps we went through earlier as follows:

$$\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 3 & 2 & -1 & 1 \\ 2 & 0 & 1 & -1 \end{bmatrix} \xrightarrow{-3R_1+R_2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 2 & 0 & 1 & -1 \end{bmatrix}$$

$$\xrightarrow{-2R_1+R_3} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 0 & -2 & -1 & -13 \end{bmatrix} \xrightarrow[-R_3]{-R_2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 2 & 1 & 13 \end{bmatrix} \xrightarrow{-2R_2+R_3} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & -7 & -21 \end{bmatrix}$$

$$\xrightarrow{-\frac{1}{7}R_2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & 1 & 3 \end{bmatrix} \xrightarrow{-4R_3+R_2} \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix} \xrightarrow{-R_3+R_1} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

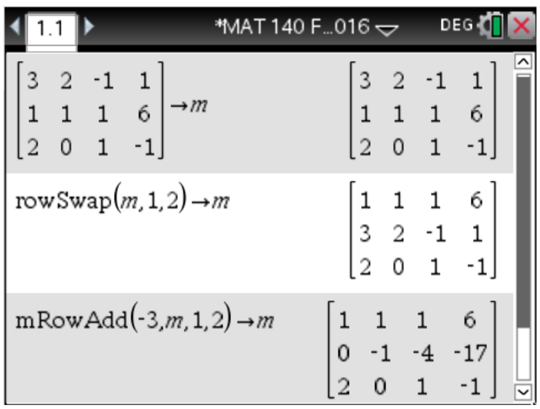
$$\xrightarrow{-R_2+R_1} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

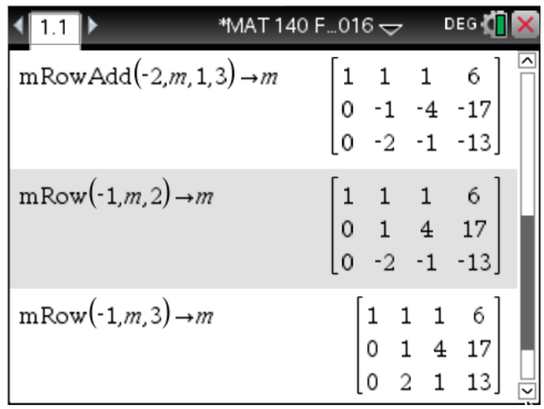
This procedure is called **row reduction**, or **Gauss-Jordan elimination**.

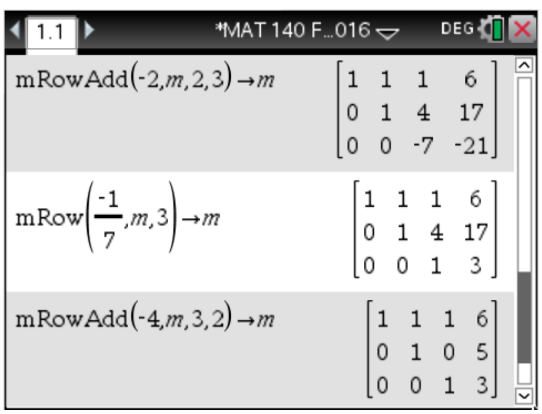
Row reduction with the TI-Nspire

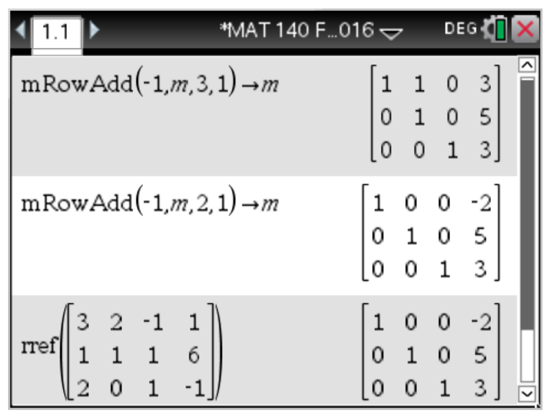
Let's do this with the TI-Nspire.

We begin by storing the matrix $\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix}$ in m. In fact, every time we do a row operation we store the new matrix back into m so that we can call m in the next operation:









It is important that you know this procedure, and how each steps works. Later it helps us to understand how it is related to solving systems of equations, which will be our focus in the next section. Of course the TI-Nspire can do all these steps in one stroke! There is a function called **rref**, which stands for 'reduced row echelon form'. This will transform a matrix into its reduced row echelon form in one stroke (as you can see in the last line of the previous calculator screen).

There is also a function **ref** which will return a 'row echelon form'—not reduced, i.e. only the first three conditions listed on page 85 are satisfied. Note the reduced form (rref) of a matrix is **unique**, but the non reduced form (ref) is **not** [and depends on the steps we take].

Row reduction. Gauss-Jordan Elimination.

Let's formalize the procedure (see also the illustrations to visualize it)

Row reduction. Input: matrix m

Phase I: (1) If $m = O$ or is empty (no rows) we are done, i.e. Phase I terminates.

(2) Move all zero rows to the bottom of the matrix, by row swaps.

(3) Select the row with the left most non zero element. If there is more than one pick the row closest to the top of the matrix. Swap this row with row 1, and make its pivot 1, using row multiplication.

(4) If there are any rows below row 1, that have a non zero element, say k, below the pivot of the first row, subtract k times row 1 from this row. Do this until there are only zeros below the pivot of row 1.

(5) Repeat Phase I on the sub matrix of m consisting of all the entries below *and* to the right of the pivot.

Phase I will terminate when a submatrix to be reduced is a zero matrix or when there are no more submatrices to reduce. **Row echelon** form has been reached i.e. properties (1), (2) and (3) of the definition of reduced row echelon form are satisfied, but maybe not yet (4). Now Phase II kicks in to produce *reduced* row echelon form.

Phase II: (1) If all entries above the pivots are zero Phase II terminates.

(2) Using the right (or left) most pivot create zeros above it. [mRowAdd]

(3) Repeat Phase II on the sub matrix containing all entries above *and* to the left (or right) of the previous pivot.

Illustrations of the steps of row reduction

Phase I (2) Move all zeros to the bottom

$$\begin{bmatrix} \\ \\ \\ \\ \boxed{\begin{matrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{matrix}} \end{bmatrix}$$

Phase I (3) Pick pivot, move its row to top, and scale pivot to 1 [mRow]

$$\begin{bmatrix} 0 & 0 & 0 & 2 & * & & * & * \\ 0 & 2 & 8 & 4 & * & & * & * \\ 0 & 0 & 3 & * & * & & * & * \\ 0 & 4 & * & * & * & & * & * \\ 0 & 7 & * & * & * & & * & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 4 & 2 & * & & * & * \\ 0 & 0 & 0 & 2 & * & & * & * \\ 0 & 0 & 3 & * & * & & * & * \\ 0 & 4 & * & * & * & & * & * \\ 0 & 7 & * & * & * & & * & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Phase I (4) create zeros below the pivot [mRowAdd]

$$\begin{bmatrix} 0 & 1 & 4 & 2 & * & & * & * \\ 0 & 0 & 0 & 2 & * & & * & * \\ 0 & 0 & 3 & * & * & & * & * \\ 0 & 0 & * & * & * & & * & * \\ 0 & 0 & * & * & * & & * & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Phase I (5) repeat Phase I for the indicated sub matrix

$$\begin{bmatrix} 0 & 1 & 4 & 2 & * & & * & * \\ 0 & 0 & 0 & 2 & * & & * & * \\ 0 & 0 & 3 & * & * & & * & * \\ 0 & 0 & * & * & * & & * & * \\ 0 & 0 & * & * & * & & * & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Phase II (2) With the right most pivot create zeros above it.

$$\begin{bmatrix} 0 & 1 & * & * & * & * & \cdots & * & * & * \\ 0 & 0 & 0 & 1 & * & * & \cdots & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & * & * & * & * & \cdots & * & 0 & * \\ 0 & 0 & 0 & 1 & * & * & \cdots & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

Phase II (3) repeat Phase II for the indicated sub matrix

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

Row reduction using elementary row matrices

Every one of the elementary row operations can be accomplished by a matrix multiplication (these matrices we call **elementary row matrices**):

Swapping row 1 and row 2:

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

Multiplying row 3 by $-\dfrac{1}{7}$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{7} \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & -7 & -21 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Adding -3 times row 1 to row 2

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 3 & 2 & -1 & 1 \\ 2 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 2 & 0 & 1 & -1 \end{bmatrix}$$

Two observations here are important:

(1) The matrix that we need we get by taking the identity matrix and performing exactly the operation we want to it.
(2) The matrices are invertible. In fact every single one of the row operations can be undone.

Swapping row 1 and row 2 is undone using the same matrix

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

Multiplying row 3 by -7 undoes multiplying row 3 by $-\frac{1}{7}$:

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

Adding 3 times row 1 to row 2 undoes adding -3 times row 1 to row 2

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

Hence we could have performed the row reduction using elementary row matrices:

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

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 3 & 2 & -1 & 1 \\ 2 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 2 & 0 & 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 2 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 0 & -2 & -1 & -13 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -1 & -4 & -17 \\ 0 & -2 & -1 & -13 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 2 & 1 & 13 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 2 & 1 & 13 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & -7 & -21 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{7} \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & -7 & -21 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 4 & 17 \\ 0 & 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

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

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

Therefore if we put all these matrix products together we get

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{7} \end{bmatrix} *$$
$$* \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} *$$
$$* \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Let's adopt the following notation for the elementary row matrices:

$S(n,m)$ is the matrix that swaps rows n and m.

$M(k;n)$ is the matrix that multiplies row n by k.

$A(k;n,m)$ is the matrix that adds k times row n to row m.

Then the previous equation can be written as

$$A(-1;2,1) * A(-1;3,2) * A(-4;3,2) * M(-\tfrac{1}{7};3) * A(-2;2,3) * M(-1;3) *$$
$$* M(-1;2) * A(-2;1,3) * A(-3;1,2) * S(1,2) * \begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Note that we need only a *finite number* of elementary row matrices E_i to produce the row reduced echelon form of a matrix, say M :

$$E_k \cdot E_{k-1} \cdot \cdots \cdot E_3 \cdot E_2 \cdot E_1 \cdot A = M$$

[Here M is the row reduced echelon form and the E_i s, the various elementary row matrices needed]. Hence there is an matrix Q that would do it all in one scoop, namely:

$$Q \;=\; E_k \cdot E_{k-1} \cdot \cdots \cdot E_3 \cdot E_2 \cdot E_1$$

Then $Q \cdot A = M$. In our case the Q is

$$Q \;=\; \frac{1}{7}\begin{bmatrix} 1 & -2 & 3 \\ 1 & 5 & -4 \\ -2 & 4 & 1 \end{bmatrix} \quad \text{and} \quad \frac{1}{7}\begin{bmatrix} 1 & -2 & 3 \\ 1 & 5 & -4 \\ -2 & 4 & 1 \end{bmatrix} \overbrace{\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix}}^{A} = \overbrace{\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}}^{M}$$

In particular since the E_i s are all invertible Q is also invertible! In fact

$$Q^{-1} \;=\; E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot \cdots \cdot E_{k-1}^{-1} \cdot E_k^{-1}$$

To demonstrate that the elementary row matrices are invertible:

(I) $S(n,m) \cdot S(n,m) = I$

$$\begin{array}{cc} & \begin{array}{cccccccccc} & & & n & & & m & & & \end{array} \\ \begin{array}{c} \\ \\ \\ n \\ \\ \\ m \\ \\ \\ \\ \end{array} & \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix} \end{array} \cdot \begin{array}{c} \begin{array}{cccccccccc} & & & n & & & m & & & \end{array} \\ \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix} \end{array} = I$$

(II) $M(k;n)\cdot M(\frac{1}{k};n)$

$$
\begin{array}{cc}
 & \begin{array}{cccccccccc} & & & n & & & & & & \end{array} \\
\begin{array}{c} \\ \\ \\ n \\ \\ \\ \\ \\ \\ \\ \end{array} &
\begin{bmatrix}
1 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & & & & & \vdots & \vdots \\
0 & 0 & \cdots & k & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \ddots & & & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & \ddots & & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & & \ddots & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & & & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{bmatrix}
\end{array}
\cdot
\begin{array}{c}
\begin{array}{cccccccccc} & & & n & & & & & & \end{array} \\
\begin{bmatrix}
1 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & & & & & \vdots & \vdots \\
0 & 0 & \cdots & \frac{1}{k} & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \ddots & & & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & \ddots & & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & & \ddots & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & & & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{bmatrix}
\end{array}
= I
$$

(III) $A(k,n,m)\cdot A(-k,n,m) \;=\; I$

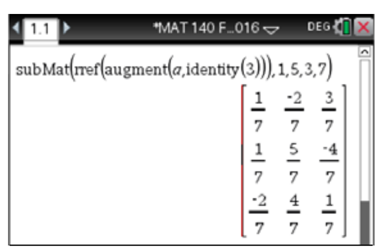
$$
\begin{array}{cc}
 & \begin{array}{cccccccccc} & & & n & & & m & & & \end{array} \\
\begin{array}{c} \\ \\ \\ n \\ \\ \\ m \\ \\ \\ \\ \end{array} &
\begin{bmatrix}
1 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & & & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 1 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \ddots & & \vdots & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & \ddots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & k & \cdots & \cdots & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & & & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 1
\end{bmatrix}
\end{array}
\cdot
\begin{array}{c}
\begin{array}{cccccccccc} & & & n & & & m & & & \end{array} \\
\begin{bmatrix}
1 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & & & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 1 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \ddots & & \vdots & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & & \ddots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & -k & \cdots & \cdots & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & & & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & \cdots & 0 & 1
\end{bmatrix}
\end{array}
= I
$$

To find Q using the TI-Nspire

If A is an $n\times m$ matrix the TI-Nspire will produce Q as follows

subMat(rref(augment(A**, identity(**n**))),** 1**,** $m+1$**,** n **,** $n+m$ **)**

$$A=\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix} \text{ then } Q \;=\; \frac{1}{7}\begin{bmatrix} 1 & -2 & 3 \\ 1 & 5 & -4 \\ -2 & 4 & 1 \end{bmatrix}$$



15. Solving Systems of Linear Equations

Consider the following three problems:

Problem 1: Find in $\mathbb{R}^3$ the intersection of the three planes $\begin{cases} 3x+2y-z &= 1 \\ \;\;x+\;y+z &= 6 \\ 2x \qquad\;\; +z &= -1 \end{cases}$,

i.e. find all points $(x, y, z) \in \mathbb{R}^3$ that satisfy these three equations.

Problem 2: Write the vector $\begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}$ as a linear combination of $\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$,

i.e. find $x, y, z \in \mathbb{R}$ such that $x\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}$

Problem 3: Find all vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ such that $\begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}$.

Even though these problems ask different questions they can all be solved with the same technique: row reduction.

Problem 1 and 2 both are equivalent to problem 3:

$$\begin{cases} 3x+2y-z &= 1 \\ \;\;x+\;y+z &= 6 \\ 2x \qquad\;\; +z &= -1 \end{cases} \Leftrightarrow \begin{bmatrix} 3x+2y-z \\ x+\;y+z \\ 2x \qquad\;\; +z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}$$

and

$$x\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 3x+2y-z \\ x+\;y+z \\ 2x \qquad\;\; +z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}$$

But as we saw in the previous section we can find a matrix Q that transforms a matrix to its row reduced echelon form:

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

[In fact we found that $Q = \frac{1}{7}\begin{bmatrix} 1 & -2 & 3 \\ 1 & 5 & -4 \\ -2 & 4 & 1 \end{bmatrix}$ and that Q is basically a product of elementary row matrices, **and** that Q is **invertible**!]

In particular from $Q \cdot \begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$ we deduce that

$$Q \cdot \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad Q \cdot \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}$$

Hence if we take

$$\begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}$$

and multiply by Q we get

$$Q \cdot \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = Q \cdot \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}$$

which is equal to

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}$$

so that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}$$

and we have solved all three problems simultaneously. [One can easily check that the point $(x, y, z) = (-2, 5, 3)$ is indeed on all three planes of problem 1, and that in problem 2

$$-2\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} + 5\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix} \quad \text{and in problem 3} \quad \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ -1 \end{bmatrix}.]$$

Suppose we had started out with the following system of equations:

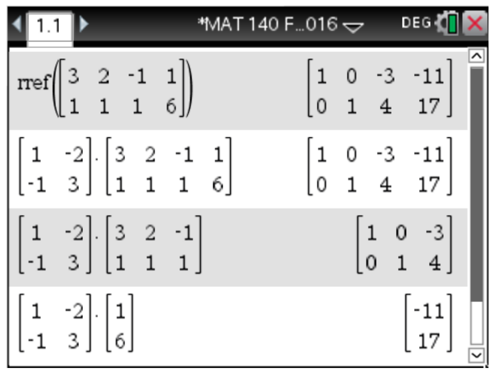
$$\begin{cases} 3x+2y-z &= 1 \\ x+\ \ y+z &= 6 \end{cases}$$

These equations represent in $\mathbb{R}^3$ two intersecting planes. The problem here is to find all points $(x, y, z) \in \mathbb{R}^3$ on the intersection, i.e. on both planes simultaneously.

Note that we can rewrite this system as $\begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$. If we then look at the

augmented matrix $\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \end{bmatrix}$ and we

row reduce it we get $\begin{bmatrix} 1 & 0 & -3 & -11 \\ 0 & 1 & 4 & 17 \end{bmatrix}$.



Hence we know there is a matrix Q such that

$$Q \cdot \begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -3 & -11 \\ 0 & 1 & 4 & 17 \end{bmatrix}$$

in fact $Q = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}$.

Hence it follows that $Q \cdot \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \end{bmatrix}$ and $Q \cdot \begin{bmatrix} 1 \\ 6 \end{bmatrix} = \begin{bmatrix} -11 \\ 17 \end{bmatrix}$

So that the original equation $\begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$ can be rewritten as follows

$$Q \cdot \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = Q \cdot \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

which give us

$$\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -11 \\ 17 \end{bmatrix}$$

so that

$$\begin{cases} x \quad -3z &= -11 \\ \quad y+4z &= 17 \end{cases}.$$

At first it might not be obvious what we have accomplished. We have replaced the system of equations

$$(1)\ \cdots \begin{cases} 3x+2y-z &= 1 \\ x+\ y+z &= 6 \end{cases}$$

with

$$(2)\ \cdots \begin{cases} x \qquad -3z &= -11 \\ \quad y+4z &= 17 \end{cases}$$

It is important to observe that both systems have **exactly** the same solution set:

(a) Any point (x_0, y_0, z_0) satisfying the system of equations (1) also satisfies the system of equations (2) (just use Q) :

$$\begin{cases} 3x_0+2y_0-z_0 &= 1 \\ x_0+\ y_0+z_0 &= 6 \end{cases} \Rightarrow \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

$$\Rightarrow \overbrace{\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}}^{Q} \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \overbrace{\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}}^{Q} \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} -11 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{cases} x_0 \qquad -3z_0 &= -11 \\ \quad y_0+4z_0 &= 17 \end{cases}$$

and

(b) any point (x_0, y_0, z_0) satisfying the system of equations (2) also satisfies the system of equations (1) (just use Q^{-1})

$$\begin{cases} x_0 \qquad -3z_0 &= -11 \\ \quad y_0+4z_0 &= 17 \end{cases} \Rightarrow \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} -11 \\ 17 \end{bmatrix}$$

$$\Rightarrow \overbrace{\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}}^{Q^{-1}} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \overbrace{\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}}^{Q^{-1}} \begin{bmatrix} -11 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 3x_0+2y_0-z_0 &= 1 \\ x_0+\ y_0+z_0 &= 6 \end{cases}$$

That is why it was so important to use *invertible* operations: we never altered the set of solutions by the operations of row reduction.

Hence the set of solutions of system (2) gives us the solutions of system (1). But even though the new system looks a little less congested it may not, at first, be apparent to the untrained eye that we are actually looking at the solution.

Recall that the initial system of equations represents two intersecting planes in $\mathbb{R}^3$. Hence we are looking for a **line** as our answer!

We can read off a solution from

$$\begin{cases} x \quad - 3z = -11 \\ \quad y + 4z = 17 \end{cases}$$

by slightly rewriting it to get

$$\begin{cases} x = -11 + 3z \\ y = \ \ 17 - 4z \end{cases}$$

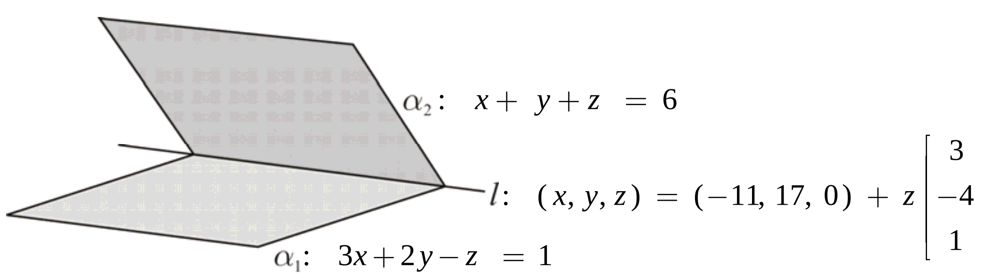
and if we throw in the obvious fact that $z = z$ we get

$$\begin{cases} x = -11 + 3z \\ y = \ \ 17 - 4z \\ z = \qquad\quad z \end{cases}$$

But this is the parametric equation of the line: $(x, y, z) = (-11, 17, 0) + z \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$

where the parameter in this case is z. This *is* the line we were looking for. This we can easily check, by 'substituting' the line into the two planes $\begin{cases} 3x + 2y - z = 1 \\ x + y + z = 6 \end{cases}$:

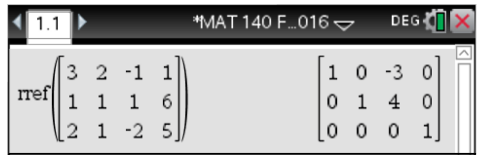
$$\begin{cases} 3(-11 + 3z) + 2(17 - 4z) - z = 1 \\ (-11 + 3z) + (17 - 4z) + z = 6 \end{cases} \Leftrightarrow \begin{cases} 1 = 1 \ \checkmark \\ 6 = 6 \ \checkmark \end{cases}$$



Let's look at another system of equations in $\mathbb{R}^3$: $\begin{cases} 3x+2y-z &= 1 \\ x+\ y+z &= 6 \\ 2x+y-2z &= 5 \end{cases}$

We have here three planes in $\mathbb{R}^3$ and are looking for their intersection, i.e. all points $(x, y, z) \in \mathbb{R}^3$ that satisfy these three equations.

Row reducing $\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 1 & -2 & 5 \end{bmatrix}$ we get $\begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$,



i.e. there is an invertible matrix Q such that

$$Q\begin{bmatrix} 3 & 2 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ 2 & 1 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad [\text{ In fact: } Q = \frac{1}{10}\begin{bmatrix} -1 & -9 & 11 \\ 7 & 13 & -17 \\ -1 & 1 & 1 \end{bmatrix}\ ]$$

and in particular

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

Hence we can transform our initial system

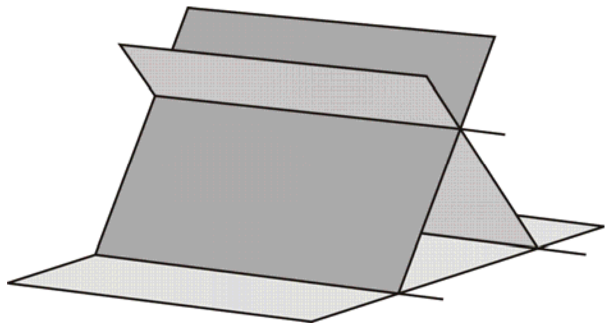
$$\begin{aligned} \begin{cases} 3x+2y-z &= 1 \\ x+\ y+z &= 6 \\ 2x+y-2z &= 5 \end{cases} &\Leftrightarrow \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix} \\ &\Leftrightarrow Q\begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = Q\begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix} \\ &\Leftrightarrow \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &\Leftrightarrow \begin{cases} x \qquad -3z &= 0 \\ y+4z &= 0 \\ 0 &= 1 \end{cases} \end{aligned}$$

Since Q is invertible we have established that the following two systems are equivalent, i.e. they have exactly the same solution set:

$$\begin{cases} 3x+2y-z &= 1 \\ x+\ y+z &= 6 \\ 2x+y-2z &= 5 \end{cases} \Leftrightarrow \begin{cases} x \qquad -3z &= 0 \\ y+4z &= 0 \\ 0 &= 1 \end{cases}$$

But clearly the system $\begin{cases} x \quad -3z = 0 \\ \quad y+4z = 0 \\ \qquad \boxed{0 = 1} \end{cases}$ has no solutions, since there are no $(x, y, z) \in \mathbb{R}^3$ that satisfy $0 = 1$. Hence the initial system has no solutions either.

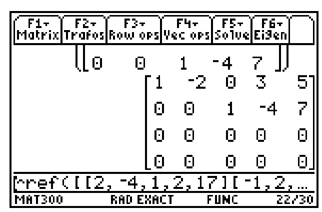
In fact the system $\begin{cases} 3x+2y-z = 1 \\ x+y+z = 6 \\ 2x+y-2z = 5 \end{cases}$ corresponds geometrically to three planes intersecting pair wise in parallel lines:



Let's look at another system of equations in this time in $\mathbb{R}^4$: $\begin{cases} 2w-4x+y+2z = 17 \\ -w+2x+y-7z = 2 \\ 3w-6x+y+5z = 22 \\ \qquad\qquad y-4z = 7 \end{cases}$

We have here four hyper-planes in $\mathbb{R}^4$ and are looking for their intersection, i.e. all points $(w, x, y, z) \in \mathbb{R}^4$ that satisfy these four equations.

The augmented matrix $\begin{bmatrix} 2 & -4 & 1 & 2 & 17 \\ -1 & 2 & 1 & -7 & 2 \\ 3 & -6 & 1 & 5 & 22 \\ 0 & 0 & 1 & -4 & 7 \end{bmatrix}$ rrefs to,



i.e. there is an invertible matrix Q such that

$$Q\begin{bmatrix} 2 & -4 & 1 & 2 & 17 \\ -1 & 2 & 1 & -7 & 2 \\ 3 & -6 & 1 & 5 & 22 \\ 0 & 0 & 1 & -4 & 7 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 3 & 5 \\ 0 & 0 & 1 & -4 & 7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and in particular

$$Q\begin{bmatrix} 2 & -4 & 1 & 2 \\ -1 & 2 & 1 & -7 \\ 3 & -6 & 1 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad Q\begin{bmatrix} 17 \\ 2 \\ 22 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 0 \\ 0 \end{bmatrix}$$

Hence our initial system is transformed by row reduction to another system:

$$\begin{cases} 2w-4x+y+2z &= 17 \\ -w+2x+y-7z &= 2 \\ 3w-6x+y+5z &= 22 \\ y-4z &= 7 \end{cases} \quad \Leftrightarrow \quad \begin{cases} w-2x \quad +3z = 5 \\ \qquad\quad y-4z = 7 \end{cases}$$

These systems have **exactly** the same solutions set, using the same reasoning as before with the *invertible* matrix Q [of course a different Q for this case]:

$$\begin{cases} 2w-4x+y+2z &= 17 \\ -w+2x+y-7z &= 2 \\ 3w-6x+y+5z &= 22 \\ y-4z &= 7 \end{cases} \Leftrightarrow \begin{bmatrix} 2 & -4 & 1 & 2 \\ -1 & 2 & 1 & -7 \\ 3 & -6 & 1 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix}\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 17 \\ 2 \\ 22 \\ 7 \end{bmatrix}$$

$$\Leftrightarrow \quad Q\begin{bmatrix} 2 & -4 & 1 & 2 \\ -1 & 2 & 1 & -7 \\ 3 & -6 & 1 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix}\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = Q\begin{bmatrix} 17 \\ 2 \\ 22 \\ 7 \end{bmatrix}$$

$$\Leftrightarrow \quad \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \quad \begin{cases} w-2x \quad +3z = 5 \\ \qquad\quad y-4z = 7 \end{cases}$$

What remains is to extract the solution from our last system of equations:

$$\begin{cases} w = 5+2x \quad -3z \\ y = 7 \qquad\quad +4z \end{cases} \Leftrightarrow \begin{cases} w = 5+2x \quad -3z \\ x = \qquad x \\ y = 7 \qquad\quad +4z \\ z = \qquad\qquad z \end{cases}$$

$$\Leftrightarrow \quad \boxed{(w,x,y,z) = (5,0,7,0) + x\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + z\begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \end{bmatrix}}$$

Hence the solution of the system consists of all points on a two dimensional plane in $\mathbb{R}^4$

through the point $(5,0,7,0)$ and with direction vectors $\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \end{bmatrix}$

Most books on Linear Algebra will actually introduce row reduction, or Gauss-Jordan elimination, in the setting of equations, rather than matrices and augmented matrices. They will discuss the operations that are allowed

Type I: Swapping two equations.
Type II: Multiplying an equation by a non zero constant.
Type III: Adding a multiple of one equation to another equation.

and proceed to solve the system of equations $\begin{cases} 3x+2y-z &= 1 \\ x+y+z &= 6 \\ 2x+z &= -1 \end{cases}$ as follows:

$$\begin{cases} 3x+2y-z = 1 \\ x+y+z = 6 \\ 2x+z = -1 \end{cases} \xrightarrow{\text{Eq}_1\leftrightarrow\text{Eq}_2} \begin{cases} x+y+z = 6 \\ 3x+2y-z = 1 \\ 2x+z = -1 \end{cases}$$

$$\xrightarrow{\text{Eq}_2-3\text{Eq}_1} \begin{cases} x+y+z = 6 \\ -y-4z = -17 \\ 2x+z = -1 \end{cases} \xrightarrow{\text{Eq}_3-2\text{Eq}_1} \begin{cases} x+y+z = 6 \\ -y-4z = -17 \\ -2y-z = -13 \end{cases}$$

$$\xrightarrow[-\text{Eq}_3]{-\text{Eq}_2} \begin{cases} x+y+z = 6 \\ y+4z = 17 \\ 2y+z = 13 \end{cases} \xrightarrow{\text{Eq}_3-2\text{Eq}_1} \begin{cases} x+y+z = 6 \\ y+4z = 17 \\ -7z = -21 \end{cases}$$

$$\xrightarrow{-\frac{1}{7}\text{Eq}_3} \begin{cases} x+y+z = 6 \\ y+4z = 17 \\ z = 3 \end{cases} \xrightarrow{\text{Eq}_2-4\text{Eq}_3} \begin{cases} x+y+z = 6 \\ y = 5 \\ z = 3 \end{cases}$$

$$\xrightarrow{\text{Eq}_1-\text{Eq}_3} \begin{cases} x+y = 3 \\ y = 5 \\ z = 3 \end{cases} \xrightarrow{\text{Eq}_1-\text{Eq}_2} \begin{cases} x = -2 \\ y = 5 \\ z = 3 \end{cases}$$

Only after such a discussion are matrices introduced.

Of course when we compare this with the way we introduced solving systems of equations and row reduction, in this and the previous section, we notice two things: (1) essentially the same thing is going on, and (2) it is easier to work with matrices than equations. [e.g. we don't have to keep writing x, y and z everywhere.]. That's why we started out with matrices and when we need to solve a system of equations we immediately translate them into matrix equations.

Some musings about the importance of Row reduction

Row reduction is at the heart of linear algebra. It is *the* tool used in linear algebra. Let's look at some of the things we are able to do with it.

(1) Intersection of two planes

In chapter 8 we discussed the intersection of two planes, and gave the following example

Example: Intersect $\alpha: \; 3x-2y-z=15$ and $\beta: \; 3x+2y+7z=3$.

$$\begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix} \times \begin{bmatrix} 3 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} -12 \\ -24 \\ 12 \end{bmatrix}$$ so we could take $\vec{n} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ as a direction vector for the line.

To find a point on the line select $z=0$ then $\begin{cases} 3x-2y=15 \\ 3x+2y=3 \end{cases}$ which gives us $\begin{cases} x=3 \\ y=-3 \end{cases}$.

Hence $(3,-3,0)$ is on both planes, so that a vector equation of the line would be

$$l: \; (x,y,z) = (3,-3,0)+t\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Here is another way of doing this: substitute the *parametric* equations of one plane into the *normal* equation of the other

Example: $\alpha: \; 3x-2y-z=15$ and $\beta: \; 3x+2y+7z=3$.

Rewrite α as $\begin{cases} x=5+t+2s \\ y=3s \\ z=3t \end{cases}$ and substitute this in $\beta: \; 3x+2y+7z=3$:

to get

$$3(5+t+2s)+2(3s)+7(3t)=3 \quad \Rightarrow \quad 12s+24t=-12 \quad \Rightarrow \quad s=-2t-1$$

hence we find

$$\begin{cases} x=5+t+2(-2t-1) \\ y=3(-2t-1) \\ z=3t \end{cases} \quad \Rightarrow \quad \begin{cases} x=3-3t \\ y=-3-6t \\ z=3t \end{cases}$$

i.e. the line $(x,y,z) = (3,-3,0)+\tilde{t}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$

Now that we have row reduction available we can do this more easily as follows:

$$\begin{cases} 3x-2y-z=15 \\ 3x+2y+7z=3 \end{cases}$$

$$\Rightarrow \begin{cases} x \quad + z = 3 \\ \quad y+2z=-3 \end{cases} \qquad \text{since rref}\begin{bmatrix} 3 & -2 & -1 & 15 \\ 3 & 2 & 7 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 2 & -3 \end{bmatrix}$$

$$\Rightarrow \begin{cases} x = 3-z \\ y=-3-2z \\ z = z \end{cases}$$

$$\Rightarrow (x,y,z) = (3,-3,0) + z\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$$

(2) Intersection of two lines

In chapter 8 we also considered the intersection of the two lines

$$l: \begin{cases} x=2+2s \\ y=1+4s \\ z=4-2s \end{cases} \quad \text{and} \quad m: \begin{cases} x = 7+r \\ y=13+4r \\ z=-3-3r \end{cases}$$

i.e.

$$\begin{cases} 2+2s = 7+r \\ 1+4s = 13+4r \\ 4-2s = -3-3r \end{cases} \Rightarrow \begin{cases} 2s-r=5 \\ s-r=3 \\ 2s-3r=7 \end{cases} \Rightarrow \begin{cases} s=2 \\ r=-1 \end{cases}$$

$$\text{since rref}\begin{bmatrix} 2 & -1 & 5 \\ 1 & -1 & 3 \\ 2 & -3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

so that the point of intersection is

$$\begin{cases} x=2+2\cdot 2 \\ y=1+4\cdot 2 \\ z=4-2\cdot 2 \end{cases} \Rightarrow (6,9,0) \qquad [\text{ Or using } r=-1: \begin{cases} x = 7+(-1) = 6 \\ y=13+4(-1) = 9 \\ z=-3-3(-1) = 0 \end{cases} .]$$

(3) Expressing a vector as a linear combination of other vectors:

Suppose we want to create $\begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}$ from three given vectors: $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$:

i.e. $x\begin{bmatrix}1\\2\\1\end{bmatrix} + y\begin{bmatrix}2\\-1\\-1\end{bmatrix} + z\begin{bmatrix}-1\\1\\1\end{bmatrix} = \begin{bmatrix}5\\3\\2\end{bmatrix}$ then, as we saw before, this can easily been done with row reduction

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

i.e. $1\cdot\begin{bmatrix}1\\2\\1\end{bmatrix} + 5\cdot\begin{bmatrix}2\\-1\\-1\end{bmatrix} + 6\cdot\begin{bmatrix}-1\\1\\1\end{bmatrix} = \begin{bmatrix}5\\3\\2\end{bmatrix}$.

(4) The cross product. We introduced the cross products 'out of the blue' as

$$\vec{u}\times\vec{v} = \begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}\times\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix} := \begin{bmatrix}u_2v_3-u_3v_2\\ u_3v_1-u_1v_3\\ u_1v_2-u_2v_1\end{bmatrix}$$

and 'discovered' that it is a vector perpendicular to both $\vec{u}$ and $\vec{v}$. In that same section we explained how we could 'derive' this algebraically. Using row reduction this may now be clearer (we'll do an example first and then the general case.)

Suppose we want a vector $\begin{bmatrix}x\\y\\z\end{bmatrix}$ such that $\begin{bmatrix}x\\y\\z\end{bmatrix}\perp\begin{bmatrix}1\\-1\\2\end{bmatrix}$ ***and*** $\begin{bmatrix}x\\y\\z\end{bmatrix}\perp\begin{bmatrix}3\\5\\1\end{bmatrix}$,

i.e. $\begin{cases}x-y+2z=0\\3x+5y+z=0\end{cases}$ then $\text{rref}\begin{bmatrix}1 & -1 & 2 & 0\\ 3 & 5 & 1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0 & \frac{11}{8} & 0\\ 0 & 1 & -\frac{5}{8} & 0\end{bmatrix}$

so that $\begin{cases}x+\frac{11}{8}z=0\\ y-\frac{5}{8}z=0\end{cases} \Rightarrow \begin{cases}x=-\frac{11}{8}z\\ y=\frac{5}{8}z\\ z=z\end{cases}$ so if we pick $z=8$ then we

find $\begin{cases}x=-11\\y=5\\z=8\end{cases}$ which is exactly the cross product $\begin{bmatrix}1\\-1\\2\end{bmatrix}\times\begin{bmatrix}3\\5\\1\end{bmatrix} = \begin{bmatrix}-11\\5\\8\end{bmatrix}$.

This works in general: Suppose we want $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \perp \vec{u}$ ***and*** $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \perp \vec{v}$,

i.e. $\begin{cases} x u_1 + y u_2 + z u_3 = 0 \\ x v_1 + y v_2 + z v_3 = 0 \end{cases}$

then $\quad \text{rref}\begin{bmatrix} u_1 & u_2 & u_3 & 0 \\ v_1 & v_2 & v_3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\dfrac{u_2 v_3 - u_3 v_2}{u_1 v_2 - u_2 v_1} & 0 \\ 0 & 1 & -\dfrac{u_3 v_1 - u_1 v_3}{u_1 v_2 - u_2 v_1} & 0 \end{bmatrix}$

so that $\begin{cases} x - \dfrac{u_2 v_3 - u_3 v_2}{u_1 v_2 - u_2 v_1} z = 0 \\ y - \dfrac{u_3 v_1 - u_1 v_3}{u_1 v_2 - u_2 v_1} z = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} x = \dfrac{u_2 v_3 - u_3 v_2}{u_1 v_2 - u_2 v_1} z \\ y = \dfrac{u_3 v_1 - u_1 v_3}{u_1 v_2 - u_2 v_1} z \\ z = z \end{cases}$ so if we pick

$z = u_1 v_2 - u_2 v_1$ then we find the vector $\begin{cases} x = u_2 v_3 - u_3 v_2 \\ y = u_3 v_1 - u_1 v_3 \\ z = u_1 v_2 - u_2 v_1 \end{cases}$ which is exactly the

cross product: $\quad \vec{u} \times \vec{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} := \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}$.

(5) Being able to ‘see’ the solution after rref-ing a system of equations now also allows us to ‘see’ the parametric (and vector) equation of a plane ‘directly’:

e.g. $\quad x + 2y - 3z = 5 \quad \Rightarrow \quad \begin{cases} x = 5 - 2y + 3z \\ y = \quad\quad y \\ z = \quad\quad\quad z \end{cases}$

$$\Rightarrow \quad (x, y, z) = (5, 0, 0) + y\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$$

(6) As indicated before, the matrix Q that produces $QA = \text{rref}(A)$ we can easily find by row reduction of the matrix with the identity matrix augmented to it:

e.g. $\text{rref}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 4 \end{bmatrix}$

if we now augment the identity and then rref this we get

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

so that $\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 4 \end{bmatrix}$ i.e. $Q = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$.

Of course this makes sense since if $QA = M = \text{rref}(A)$ then

$$Q\left[\, A \mid I \,\right] = \left[\, QA \mid QI \,\right] = \left[\, M \mid Q \,\right]$$

so that: $\text{rref}\left[\, A \mid I \,\right] = \left[\, M \mid Q \,\right]$.

Here is another example:

$$\text{rref}\begin{bmatrix} 1 & -1 & 5 & 1 & 0 & 0 \\ 2 & -4 & 12 & 0 & 1 & 0 \\ -1 & 3 & -7 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 4 & 0 & \frac{3}{2} & 2 \\ 0 & 1 & -1 & 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{bmatrix}$$

so that $\begin{bmatrix} 0 & \frac{3}{2} & 2 \\ 0 & \frac{1}{2} & 1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 5 \\ 2 & -4 & 12 \\ -1 & 3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

(7) This also allows for easy computation of inverses:

If a matrix A is invertible then $\text{rref}(A) = I$. That means our Q happens to be the inverse of A! [$QA = I$].

Hence: If A is **invertible** then $\text{rref}\left[\, A \mid I \,\right] = \left[\, I \mid A^{-1} \,\right]$.

Example: $\text{rref}\left[\begin{array}{cc|cc} \underbrace{\begin{matrix} 3 & 2 \\ 8 & 5 \end{matrix}}_{A} & \underbrace{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}}_{I} \end{array}\right] = \left[\begin{array}{cc|cc} \underbrace{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}}_{I} & \underbrace{\begin{matrix} -5 & 2 \\ 8 & -3 \end{matrix}}_{A^{-1}} \end{array}\right]$

An easy check: $\begin{bmatrix} -5 & 2 \\ 8 & -3 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix}\begin{bmatrix} -5 & 2 \\ 8 & -3 \end{bmatrix}$

Another example: $\text{rref}\left[\underbrace{\begin{matrix} 1 & 3 & 1 \\ 2 & 4 & -1 \\ 1 & 1 & 1 \end{matrix}}_{A} \left| \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{I} \right.\right] = \left[\underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{I} \left| \underbrace{\begin{matrix} -\frac{5}{6} & \frac{1}{3} & \frac{7}{6} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{matrix}}_{A^{-1}} \right.\right]$

Check: $\begin{bmatrix} 1 & 3 & 1 \\ 2 & 4 & -1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} -\frac{5}{6} & \frac{1}{3} & \frac{7}{6} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{5}{6} & \frac{1}{3} & \frac{7}{6} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 1 & 3 & 1 \\ 2 & 4 & -1 \\ 1 & 1 & 1 \end{bmatrix}$

Of course when the matrix is **not invertible** we also discover that with row reduction:

e.g. $\text{rref}\left[\begin{matrix} 1 & -1 & 5 \\ 2 & -4 & 12 \\ -1 & 3 & -7 \end{matrix}\;\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}\right] = \left[\begin{matrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ \end{matrix} \atop \underbrace{\begin{matrix} 0 & 0 & 0 \end{matrix}}_{\textbf{not } I} \;\begin{matrix} 0 & \frac{3}{2} & 2 \\ 0 & \frac{1}{2} & 1 \\ 1 & -1 & 1 \end{matrix}\right]$

Hence: $\begin{bmatrix} 1 & -1 & 5 \\ 2 & -4 & 12 \\ -1 & 3 & -7 \end{bmatrix}$ is **not** invertible.

(8) Finally we will give an example how row reduction can be used to compute determinants. Of course there are many other applications of row reduction, but we will finish with this one.

Note that (1) swapping of rows creates the opposite determinant , for example

$$\det\begin{bmatrix} 1 & 3 & -2 & 6 \\ 4 & -6 & 8 & 9 \\ 0 & 7 & 5 & -4 \\ -8 & 2 & 0 & 3 \end{bmatrix} = -\det\begin{bmatrix} 1 & 3 & -2 & 6 \\ -8 & 2 & 0 & 3 \\ 0 & 7 & 5 & -4 \\ 4 & -6 & 8 & 9 \end{bmatrix}$$

(2) Adding a row to another row does not change the value of the determinant:

$$\det\begin{bmatrix} 1 & 3 & -2 & 6 \\ 4 & -6 & 8 & 9 \\ 0 & 7 & 5 & -4 \\ -8 & 2 & 0 & 3 \end{bmatrix} \xrightarrow{2R_2 + R_4} \det\begin{bmatrix} 1 & 3 & -2 & 6 \\ 4 & -6 & 8 & 9 \\ 0 & 7 & 5 & -4 \\ 0 & -10 & 16 & 21 \end{bmatrix}$$

(3) Multiplying a row with a constant k makes the determinant k times bigger:

$$\det\begin{bmatrix} 1 & 3 & -2 & 6 \\ 3\cdot 4 & 3\cdot(-6) & 3\cdot 8 & 3\cdot 9 \\ 0 & 7 & 5 & -4 \\ -8 & 2 & 0 & 3 \end{bmatrix} = 3\cdot\det\begin{bmatrix} 1 & 3 & -2 & 6 \\ 4 & -6 & 8 & 9 \\ 0 & 7 & 5 & -4 \\ 0 & -10 & 16 & 21 \end{bmatrix}$$

Hence if we keep track of these changes while we do a row reduction to produce the ref of a matrix (we don't need the rref !) we are actually computing its determinant:

Example:

$$\det\begin{bmatrix} -1 & 2 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix} = -\det\begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & 3 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{after swapping row 1 and row 2}$$

$$= -\det\begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & 4 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{after adding row 1 to row 2}$$

$$= -4\cdot\det\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{after taking a 4 out of row 2}$$

$$= -4\cdot\det\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & 2 \end{bmatrix} \quad \text{after adding } -\text{row 1 to row 3}$$

$$= -4\cdot\det\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{after adding row 2 to row 3}$$

$$= -4\cdot 3\cdot\det\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{after taking a 3 out of row 3}$$

$$= -4\cdot 3\cdot 1 = \boxed{-12} \quad \text{the last upper triangular matrix had determinant 1}$$