WHAT IS A MATRIX?

David F. Gleich

August 21, 2023

Learning objectives
1. Realize the difference between a table of data and a matrix.
2. Translate simple problems into matrix equations (such as least squares).
3. Review our notation for matrices, vectors, scalars, etc.

Let's start with the elephant in the room.

What is a matrix?

What are the things that we are going to learn to compute with in this class. Presumably you are taking it because you have some type of interest in using matrices.

Ask any American someone on the street of about the same age as I am "What is a matrix" and they are likely to look just a bit confused. They are probably thinking you meant to ask "What is The Matrix", in reference to the 1999-ish movie staring Keanu Reeves.

Now, ask someone in Purdue's college of engineering or science "What is a matrix' ' and you'll probably get some of the answers that we'll discuss below.

- A matrix is a 2-dimensional array, or table, of numbers.
- A matrix is a linear transformation.
- A matrix is a coordinate representation of a linear transformation between vector spaces.
- A matrix is an element of $\mathbb{R}^{m,n}$ or $\mathbb{C}^{m,n}$, or some other field such as $\mathbf{F}_2$, which arises frequently in cryptography.
- It's a table of numbers where linear algebraic manipulations make sense.
- I know it when I see it.
- a matrix stores vectors
- something used to simplify mathematical calculations

1 A STARTING EXAMPLE

Is the following a matrix?

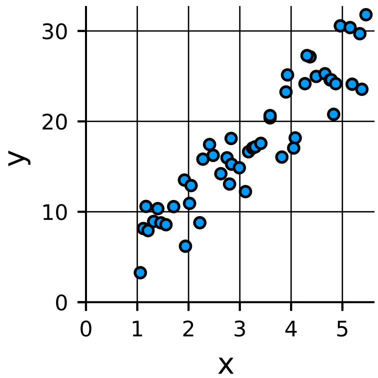
$$\begin{bmatrix} 4 & 9 & 6 & 0 & 4 & 5 & 3 \\ 4 & 9 & 4 & 1 & 9 & 9 & 7 \\ 4 & 9 & 4 & 6 & 0 & 1 & 0 \\ 4 & 9 & 6 & 1 & 7 & 6 & 1 \\ 4 & 9 & 4 & 8 & 7 & 9 & 8 \\ 4 & 9 & 6 & 2 & 3 & 9 & 9 \\ 4 & 9 & 4 & 6 & 0 & 1 & 3 \\ 4 & 9 & 6 & 9 & 4 & 3 & 2 \\ 4 & 9 & 4 & 9 & 0 & 2 & 5 \\ 4 & 9 & 4 & 6 & 0 & 0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

No! This is a table of telephone numbers, expanded by digit into columns, not a matrix.

2 AN EXAMPLE

Let's try and answer this question a different perspective and see a few places where matrices arise. This will also help introduce us to our notation in class.

Suppose we have some data and we'd like to fit a linear model to it.



x	y
1.06	3.26
1.17	10.59
1.12	8.15
1.21	7.92
1.32	8.93
1.4	10.35
1.46	8.78
1.56	8.57
1.71	10.57
1.92	13.51
1.94	6.19
2.02	10.93
2.05	12.9
2.28	15.82
2.22	8.79
2.41	17.44
2.48	16.24
2.63	14.21
2.75	15.97
2.8	13.07
2.83	18.11
2.84	15.26
2.99	14.87
3.17	16.64
3.11	12.23
3.25	17.05
3.3	17.18
3.41	17.57
3.59	20.39
3.59	20.64
3.82	16.05
3.93	25.14
3.9	23.25
4.08	18.17
4.05	17.05
4.27	24.17
4.37	27.14
4.31	27.28
4.49	24.97
4.66	25.27
4.76	24.59
4.83	20.78
4.78	24.61
4.87	24.17
4.96	30.57
5.19	24.11
5.15	30.38
5.38	23.55
5.34	29.7
5.46	31.8

The data for our problem are pairs: $(y_1, x_1), \dots, (y_N, x_N)$.[1] We can assemble them into a first matrix:

$$X = \begin{bmatrix} y_1 & x_1 \\ y_2 & x_2 \\ \vdots & \\ y_N & x_N \end{bmatrix} \text{ or } X = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \\ x_N & y_N \end{bmatrix} \text{ or } X = \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ y_1 & y_2 & \dots & y_N \end{bmatrix} \text{ or } \dots$$

Are these matrices? Let's keep going before answering this!

Our goal is to find coefficients (c_1, c_2) such that $y = c_2 x + c_1$ is a good fit to the data. There are a few ways that we could measure fit. We will be expendient and insist that for each data point, we want:

$$(y_i - (c_2 x_i + c_1))^2$$

to be as small as possible for all datapoints.[2] This gives us a way to measure how good c_1, c_2 are. Our goal is now to minimize the function:

$$f(c_1, c_2) = \sum_{i=1}^{N} (y_i - (c_2 x_i + c_1))^2.$$

Let us translate this problem with notation to eliminate the indices.

Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{c}$ be the vectors:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

then

$$f(\mathbf{c}) = \|\mathbf{y} - \mathbf{x}c_2 - \mathbf{e}c_1\|^2.$$

Here, the vector $\mathbf{e}$ is just a vector of all ones, and $\|\cdot\|$ is the 2-norm, or Euclidean norm, of a vector:

$$\|\mathbf{z}\| = \sqrt{\sum_{i=1}^{n} |z_i|^2}.$$

Now, let $\boldsymbol{A}$ be the matrix:

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{e} & \mathbf{x} \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_N \end{bmatrix}.$$

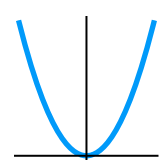
Then we can write our function f as:

$$f(\mathbf{c}) = \|\mathbf{y} - \boldsymbol{A}\mathbf{c}\|^2.$$

This type of problem is an instance of what is called a *least squares* problem. The data to the problem are encoded into a matrix $\boldsymbol{A}$ and the goal is to produce coefficients $\mathbf{c}$ that constitute a linear relationship.

[1]

[2] This is a general instances of a squared loss approximation. If we want two values such that $a \approx b$ then with squared loss, we want $(a - b)^2$. These loss functions are often illustrated with a plot:



EXAMPLE 1 *Hint that there is a matrix. Note that we can reparameterize the line as:*

$$y = d_2(x-1) + d_1$$

in which case we have

$$c_2 = d_2, c_1 = d_1 - d_2$$

or

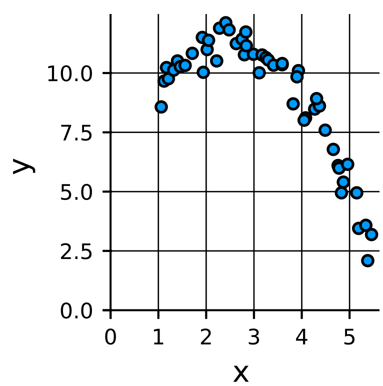
$$\mathbf{c} = \underbrace{\begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}}_{=\boldsymbol{T}}\mathbf{d}$$

which where $\boldsymbol{T}$ is definitely a matrix! In this case, we'd have:

$$\|\mathbf{y} - \boldsymbol{A}\boldsymbol{T}\mathbf{d}\|$$

if we wanted to parameterize in terms of $\boldsymbol{T}$.

Suppose instead we had the data in the following picture:



Then it's clear from the picture that a linear model isn't appropriate. We can still solve the problem we just posed, but there may be other models of our data that are appropriate. Such as a quadratic:

$$c_3x^2 + c_2x + c_1.$$

By going through the same type of steps, we can write the resulting problem in terms of the matrix:

$$\boldsymbol{A} = \begin{bmatrix}1 & x_1 & x_1^2\\ 1 & x_2 & x_2^2\\ \vdots & & \\ 1 & x_N & x_N^2\end{bmatrix}.$$

The final problem is then to find $\mathbf{c}$ to minimize:

$$f(\mathbf{c}) = \|\mathbf{y} - \boldsymbol{A}\mathbf{c}\|^2.$$

3 SUMMARY OF LEAST SQUARES EXAMPLES

In this case, we've translated *two* distinct problems into the same general form:

find $\mathbf{c}$ to make $\|\mathbf{y} - \boldsymbol{A}\mathbf{c}\|^2$ as small as possible.

This idea underlies our class. There are a great many problems in science, engineering, biology, and everywhere that can be turned into common matrix problems.

Here, the data are.

x	y
1.06	8.57
1.17	10.23
1.12	9.66
1.21	9.76
1.32	10.12
1.4	10.51
1.46	10.27
1.56	10.32
1.71	10.83
1.92	11.5
1.94	10.03
2.02	10.99
2.05	11.38
2.28	11.89
2.22	10.51
2.41	12.12
2.48	11.82
2.63	11.24
2.75	11.43
2.8	10.77
2.83	11.73
2.84	11.15
2.99	10.79
3.17	10.76
3.11	10.01
3.25	10.65
3.3	10.55
3.41	10.33
3.59	10.35
3.59	10.4
3.82	8.7
3.93	10.1
3.9	9.84
4.08	8.11
4.05	8.01
4.27	8.48
4.37	8.61
4.31	8.92
4.49	7.59
4.66	6.78
4.76	6.1
4.83	4.95
4.78	5.99
4.87	5.4
4.96	6.15
5.19	3.45
5.15	4.95
5.38	2.09
5.34	3.58
5.46	3.19

4 COMMON MATRIX PROBLEMS

The three canonical and most common matrix problems are:

- Linear systems of equations: find $\mathbf{x}$ such that $\boldsymbol{A}\mathbf{x} = \mathbf{b}$
- Least squares problems: find $\mathbf{x}$ to minimize $\|\boldsymbol{A}\mathbf{x} - \mathbf{b}\|^2$
- Eigenvalue problems: find $\mathbf{x}, \lambda$ such that $\boldsymbol{A}\mathbf{x} = \lambda\mathbf{x}$

These problems have deep relationships.

4.1 SOLVING LEAST SQUARES VIA A LINEAR SYSTEM.

We can turn the least squares problem

$$\text{find } \mathbf{x} \text{ to minimize } \|\boldsymbol{A}\mathbf{x} - \mathbf{b}\|^2$$

into a related linear system.[3]

[3] This is not necessarily the best way to solve a least squares problem. We'll see better ways in the future!

The first thing we need to do is understand how to find the minimum point of the least squares problem. There are a few ways to do this. One of the easiest is to think back to calculus class and about how we can find the extreme points of a simple quadratic function. Let $s(x) = 1/2ax^2 + b * x + c$ be a simple quadratic. This function only has a single extreme point when $s(x)$ is a minimum. Then, via calculus, we can find the extreme points by finding places where the derivative is zero. Here, $s'(x) = ax + b = 0$ gives the minimum point $x = -b/a$.

For least squares, we have :

$$f(\mathbf{x}) = \|\boldsymbol{A}\mathbf{x} - \mathbf{b}\|^2 = (\boldsymbol{A}\mathbf{x} - \mathbf{b})^T(\boldsymbol{A}\mathbf{x} - \mathbf{b})$$

This idea that we can find the extreme points by looking for points where the derivative is zero generalizes to multivariate functions such as our $f(\mathbf{x})$ for least squares. This is because $f(\mathbf{x})$ for least squares is a smooth, convex functions. That means that $f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \le \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y})$ when $0 \le \alpha \le 1$.

EXAMPLE 2 *Showing that this property $f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \le \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y})$ is a good exercise in working with vectors. The key step is to show:*

$$\alpha^2\mathbf{x}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{x} + 2\alpha(1 - \alpha)\mathbf{x}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{y} + (1 - \alpha)^2\mathbf{y}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{y} \le \alpha\mathbf{x}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{x} + (1 - \alpha)\mathbf{y}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{y}$$

This can be done via by showing that difference is less than 0.

$$\begin{aligned} LHS - RHS &= \alpha(\alpha - 1)\mathbf{x}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{x} + \alpha(\alpha - 1)\mathbf{y}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{y} - 2\alpha(\alpha - 1)\mathbf{x}^T\boldsymbol{A}^T\boldsymbol{A}\mathbf{y} \\ &= \alpha(\alpha - 1)(\mathbf{x} - \mathbf{y})^T\boldsymbol{A}^T\boldsymbol{A}(\mathbf{x} - \mathbf{y}) \\ &= \alpha(\alpha - 1)\|\boldsymbol{A}(\mathbf{x} - \mathbf{y})\| \\ &\le 0. \end{aligned}$$

The derivate or *gradient* of $f(\mathbf{x})$ is

$$f'(\mathbf{x}) = 2\boldsymbol{A}^T\boldsymbol{A}\mathbf{x} - 2\boldsymbol{A}^T\mathbf{b}.$$

This is zero when

$$mA^T\boldsymbol{A}\mathbf{x} = \boldsymbol{A}^T\mathbf{b}.$$

The result is a linear system of equations that *solve* a least squares problem.

4.2 FROM AN EIGENVECTOR PROBLEM TO A LINEAR SYSTEM

Suppose that we know that λ is an eigenvalue of $\mathbf{x}$, that means, we may want to find the eigenvector. This can be done by solving the linear systems of equations:

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\mathbf{x} = 0.$$

This isn't the standard type of linear system of equations, but it is a valid problem!

So to get back to the starting question: what is a matrix?

For our purposes:

a matrix is a table of 2d numbers where linear algebraic operations make sense on the rows or columns

So solving a least-squares problem with phone numbers doesn't make any sense. But solving this where the data come from experiments makes a good deal of sense.

The reason that we study the subject of matrix computations, and indeed, the reason that this subject continues to be interesting is that our goal is to use the *structure* of the problem in order to solve the underlying problem better.

In this case, better may mean any of these:

$$\{\text{faster, more accurately, more reliably}, \ldots .\}$$

We will see this soon!