Lab 1b: Recovering matrix from data

MATH 5110: Applied Linear Algebra and Matrix Analysis

Northeastern University
Author: Chris King

1 Introduction

The goal of this Lab is to learn how to construct a matrix which will map a given set of input vectors to a given set of output vectors.

1.1 Modeling with a linear map

We are given a 'black box' that takes an input vector $\vec{x}$ and produces an output vector $\vec{b}$. We want to model the black box using a linear map L, so the linear map should satisfy

$$L(\vec{x}) = \vec{b} \tag{1}$$

Of course if the black box is actually nonlinear then this cannot work for all pairs of input/output vectors. However if we limit the number of pairs of vectors $(\vec{x}, \vec{b})$ for which we want the relation (1) to hold, then it is possible to find such a map L even in the case where the black box is nonlinear.

To be precise about this, suppose that both the input and output vectors for the black box are n-dimensional vectors for some integer n. That is $\vec{x} \in \mathbb{R}^n$ and $\vec{b} \in \mathbb{R}^n$. Then L should be a linear map from $\mathbb{R}^n$ to $\mathbb{R}^n$. Every linear map from $\mathbb{R}^n$ to itself is equal to multiplication by a $n \times n$ matrix, and we want to find this matrix.

In order to handle the problem that the black box may not be linear, we limit the number of pairs of input/output vectors to be exactly n. That is we assume that we are given n input vectors $\vec{x}^{(1)}, \dots, \vec{x}^{(n)}$ and n output vectors $\vec{b}^{(1)}, \dots, \vec{b}^{(n)}$ such that for every $k = 1, \dots, n$ we have

$$L(\vec{x}^{(k)}) = \vec{b}^{(k)} \tag{2}$$

Let us write A to denote the $n \times n$ matrix which implements the linear map A. So we want to find the matrix A such that

$$L(\vec{x}^{(k)}) = A\,\vec{x}^{(k)} = \vec{b}^{(k)} \quad \text{for } k = 1, \dots, n \tag{3}$$

Counting the number of parameters shows that this should be possible: there are n vector equations in (3), and hence n^2 equations in total. The matrix A has n^2 entries, and these are the unknown values that we want to find. So we have n^2 equations with n^2 unknowns.

1.2 Special case $n = 2$

Let's focus for now on the case $n = 2$ in order to keep the notation simpler. And let's write the two input vectors as $\vec{x}, \vec{y}$ and the output vectors as $\vec{b}, \vec{c}$. So we have the two vector equations

$$A\,\vec{x} = \vec{b}, \quad A\,\vec{y} = \vec{c} \tag{4}$$

We will write these equations more explicitly. First we have

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{5}$$

and also

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{6}$$

Then the four equations (4) are

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 & = & b_1 \\ a_{21}x_1 + a_{22}x_2 & = & b_2 \\ a_{11}y_1 + a_{12}y_2 & = & c_1 \\ a_{21}y_1 + a_{22}y_2 & = & c_2 \end{array} \tag{7}$$

We rewrite the first and third equations as a matrix equation:

$$\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} = \begin{pmatrix} b_1 \\ c_1 \end{pmatrix} \tag{8}$$

We define

$$M = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}, \quad C_1 = \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix}, \quad D_1 = \begin{pmatrix} b_1 \\ c_1 \end{pmatrix} \tag{9}$$

so that the equation can be written

$$MC_1 = D_1 \tag{10}$$

Assuming that $\mathrm{rank}(M) = 2$, there is a unique vector C_1 which is the solution of (10). From (9) we see that this vector C_1 gives us the first row of the unknown matrix A.

We apply similar reasoning with the second and fourth equations in (7): define

$$C_2 = \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix}, \quad D_2 = \begin{pmatrix} b_2 \\ c_2 \end{pmatrix} \tag{11}$$

then we have

$$MC_2 = D_2 \tag{12}$$

Note that the same matrix M appears again. Therefore we can solve for C_2 which gives us the row of A.

To summarize: by solving (10) and (12) we get the column vectors C_1 and C_2. By taking their transposes we get the rows of A, so we have

$$A = \begin{pmatrix} C_1^T \\ C_2^T \end{pmatrix} \tag{13}$$

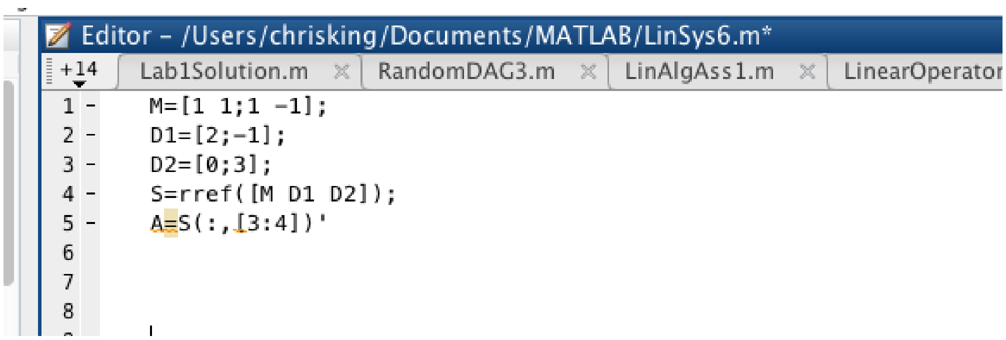
1.3 Example for $n = 2$

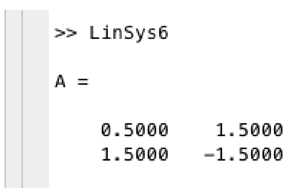
Let's see a concrete example. Suppose that

$$\vec{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \tag{14}$$

Then we get

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad D_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 0 \\ 3 \end{pmatrix} \tag{15}$$





We can check that this is correct, namely that

$$A\vec{x} = A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 3 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \vec{b}, \tag{16}$$

and similarly for $A\vec{y}$.

1.4 Independence of input and output vectors

The existence of the solutions C_1 and C_2 depended on M being nonsingular, that is having rank equal to 2. This is equivalent to the independence of its columns, and also equivalent to the independence of its rows. The rows of M are the vectors $\vec{x}$ and $\vec{y}$. So the condition that we can solve the system of equations to find A is the condition that the input vectors are independent. This condition must be satisfied for all n in the general case.

What happens if the vectors $\vec{b}$ and $\vec{c}$ are dependent? Then the vectors C_1 and C_2 will also be dependent, so the matrix A will be singular, and its rank will be less than two.

1.5 The setup for general dimension n

We follow the analysis given above and extend to the case of general dimension n. So now we have n input and n output vectors, and we want to find the $n \times n$ matrix A. The main new ingredient is developing adequate notation. First define the input vectors as

$$\vec{x}^{(1)}, \ldots \vec{x}^{(n)} \in \mathbb{R}^n \tag{17}$$

and similarly the output vectors

$$\vec{b}^{(1)}, \dots \vec{b}^{(n)} \in \mathbb{R}^n \tag{18}$$

The equations we need to solve are

$$A\vec{x}^{(k)} = \vec{b}^{(k)}, \quad k = 1, \dots, n \tag{19}$$

Proceeding as in the case $n = 2$, we rewrite these as another set of equations where the unknown vectors are the rows of A. The input vectors will again be combined into the matrix M. However recall that the *rows* of M in (9) were the input vectors. The same thing happens now, and the n input vectors are put in the rows of our $n \times n$ matrix M:

$$M = \begin{pmatrix} (\vec{x}^{(1)})^T \\ (\vec{x}^{(2)})^T \\ \vdots \\ (\vec{x}^{(n)})^T \end{pmatrix} \tag{20}$$

Note that while $\vec{x}^{(1)}$ is a column vector ($n \times 1$), its transpose $(\vec{x}^{(1)})^T$ is a row vector ($1 \times n$).

Next we see how the output vectors enter the story. To set up the notation we write the components of each vector as

$$\vec{b}^{(k)} = \begin{pmatrix} b_1^{(k)} \\ b_2^{(k)} \\ \vdots \\ b_n^{(k)} \end{pmatrix}, \quad k = 1, \dots, n \tag{21}$$

Then we define the vector

$$D_1 = \begin{pmatrix} b_1^{(1)} \\ b_1^{(2)} \\ \vdots \\ b_1^{(n)} \end{pmatrix} \tag{22}$$

Notice that we have taken the *first* component of each vector and combined these all together to make the column vector D_1. We proceed similarly with D_2, using the second components of all the output vectors, and so on. So we have

$$D_k = \begin{pmatrix} b_k^{(1)} \\ b_k^{(2)} \\ \vdots \\ b_k^{(n)} \end{pmatrix} \quad \text{for all } k = 1, \dots, n \tag{23}$$

Notice again that the kth entries of the output vectors are combined to create D_k. Then we consider the n vector equations

$$MC_k = D_k, \quad k = 1, \dots, n \tag{24}$$

The solutions of these equations give the rows of A, and we get

$$A = \begin{pmatrix} C_1^T \\ C_2^T \\ \vdots \\ C_n^T \end{pmatrix} \tag{25}$$

2 Data example

The data is drawn from measurements of PM2.5 pollution in Beijing over several years in the early twenty-first century. [Source: the UCI archive. https://archive.ics.uci.edu/ml/index.php]

We use 8 data characteristics measured at 4 time points. The characteristics are:

- Month (represented by integer in $\{1, \ldots, 12\}$)
- Time of Day (represented by hour in $[0, 24]$)
- Temperature (degrees Celcius)
- Relative humidity (percentage)
- CO concentration (mg per cubic m)
- benzene concentration (μg per cubic m)
- nitrogen oxide (parts per billion)
- nitogen dioxide (μg per cubic m)

and the data values are

Characteristic	*Time* 1	*Time* 2	*Time* 3	*Time* 4
Month	3	5	7	9
TOD	0	6	12	18
Temp	11.3	14.5	36.9	25.3
Rel.Hum.	56.8	78.3	17.2	33
CO	1.2	1.3	1.5	5.6
C6H6	3.6	6.9	8.3	31
NOX	62	108	78	578
NO2	77	81	94	204

2.1 First task

Your first task is to use this data to build a linear model whose inputs are the (Month, TOD, Temp, Rel Hum) for the 4 times, and whose ouputs are the coresponding values of (CO, C6H6, NOX, NO2). So you should find the 4×4 matrix which implements the linear map from these 4 input characteristics to the 4 output characteristics.

2.2 Second task

There is missing data for a fifth time, see below. Use the linear model to fill in the missing first four items.

Characteristic	*Time* 1	*Time* 2	*Time* 3	*Time* 4	*Time* 5
Month	3	5	7	9	
TOD	0	6	12	18	
Temp	11.3	14.5	36.9	25.3	
Rel.Hum.	56.8	78.3	17.2	33	
CO	1.2	1.3	1.5	5.6	1.7
C6H6	3.6	6.9	8.3	31	9.3
NOX	62	108	78	578	62
NO2	77	81	94	204	66