FCS worksheet 2: Matrices

Aims: The aim of this worksheet is to reinforce the ideas of the matrices lectures and in particular show how matrices can be used to simplify network operations. By the end you should be able to:

1. Add and multiply matrices
2. Use a matrix inverse to solve sets of linear equations
3. Transform inputs to a network into outputs using matrix-vector manipulations

Questions

1. If C is a 12×8 matrix, how many rows and columns does C have? If its elements are c_{ij}, write expressions for the element of C in the 4'th row and 6'th column and the elements in each of the corners of C (i.e. write them as $c_{??}$).

2. Calculate the matrix A in the following:

   (a)
   $$A = \begin{pmatrix} 4 & 6 \\ 8 & -1 \end{pmatrix} + 3$$

   (b)
   $$A = \begin{pmatrix} 8 & 2 \\ -4 & 4 \end{pmatrix} + \begin{pmatrix} 2 & -2 \\ 4 & 5 \end{pmatrix}$$

   (c)
   $$A = 3 \begin{pmatrix} 1 & 5 \\ 4 & 3 \end{pmatrix}$$

3. Multiply the following matrices:

   (a)
   $$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 8 & 2 \\ -1 & 4 \end{pmatrix}$$

   (b)
   $$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

   What do you notice about the answers? Why is this so?

4. $\underline{x}$ is the 3D (column) vector:
   $$\underline{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

   State whether the following multiplications are possible and if so, give the result: (a) $\underline{x}\,\underline{x}$, (b) $\underline{x}\,\underline{x}^T$, (c) $\underline{x}^T\underline{x}$, (d) $\underline{x}^T\underline{x}^T$.

5. Multiply the following matrices stating the dimensions of each matrix and the resultant matrix:

   (a)
   $$\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$$

   What is special about the resultant matrix (it should be diagonal)? What can you deduce about the relationship between the two matrices that were multiplied together from this result?

(b)

$$\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

What do you notice about the number of rows and columns of the resultant matrices compared to the dimensions of the matrices that were multiplied together? From this, deduce the dimensions of $C = AB$ if A is $m \times n$ and B is $n \times p$.

(c) Using the answer to question 5b, write the following equation

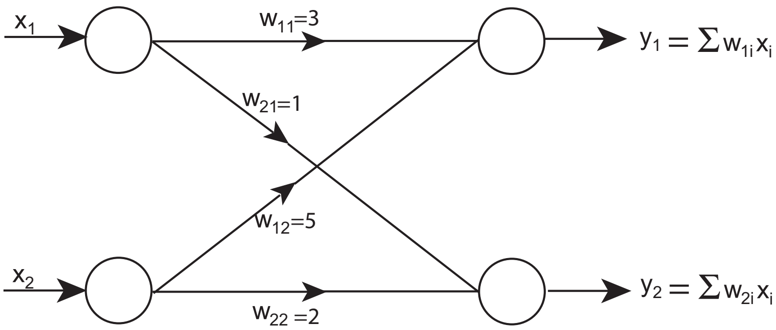
$$\begin{array}{rcrcr} y_1 & = & 3x_1 & + & 5x_2 \\ y_2 & = & x_1 & + & 2x_2 \end{array}$$

in matrix notation ie in the form:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = M \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Now re-write the matrix version of the equations using vectors $\underline{y} = (y_1, y_2)^T$ and $\underline{x} = (x_1, x_2)^T$. Using the inverse of M, M^{-1} solve the equations for $\underline{y} = (2, 2)^T$ and check your solutions (use the answer to question 5a and remember that $M^{-1}M = I$ and for any matrix A, $IA = AI = A$). NB: Here and subsequently, when I write vectors like: $\underline{y} = (y_1, y_2)^T$ it is purely for convenience. $\underline{y}$ is really a column vector, but to write it as such ruins the type-setting and takes up too much room.

6. For the linear network pictured below, write the network outputs y_1 and y_2 as weighted sums of the inputs. First re-write these equations in matrix notation, and then write them succinctly using vectors $\underline{x}$ and $\underline{y}$ for the inputs and outputs, respectively, and W as a matrix of weights, specifying the elements of W (Hint: if stuck look at questions 5b and 5c).



Now suppose we have a network with 3-dimensional input-data $\underline{x} = (x_1, x_2, x_3)^T$ and 2-dimensional output data $\underline{y} = (y_1, y_2)^T$. What are the dimensions of the weight matrix? Does the network output equation written as matrix vector multiplication change? When training the network using backpropagation, at each pass through the algorithm every weight in the network is updated using:

$$\triangle w_{ji} = -2\eta x_i \left(y_j - t_j\right)$$

where t_j is the j'th element of the 2-dimensional target vector $\underline{t} = (t_1, t_2)^T$ and $\triangle w_{ji}$ is the change in weight w_{ji}. For elegance (and ease of programming backpropagation in matlab - next week's task), we would like to be able to write the algorithm in matrix notation as: $W_{new} = W_{old} - 2\eta D$ where W_{new} and W_{old} are the new and old weight matrices respectively, η is a constant (the learning rate) and D is the matrix of weight changes, whose elements are $d_{ji} = x_i \left(y_j - t_j\right)$. To do this we need to write the equation calculating D using the vectors $\underline{x}$, $\underline{t}$ and $\underline{y}$ via matrix arithmetic.

(Hints: First make a new vector $\underline{e} = \underline{y} - \underline{t}$ and multiply $\underline{e}$ and $\underline{x}$ together. Look at question 4: there are only a limited number of ways 2 vectors can be multiplied together. For those that are possible, multiply the vectors together, calculate the results for a few elements (eg the 1st column) and see if the subscripts of the x's and y's match what the equation says they should be. Also think about what dimensions D needs to have for the equations to 'work').

Does this equation change if the dimensions of input or output vectors change?

Marks out of 10: 1 mark for questions 1-4, 3 marks for question 5 and 3 for question 6.