MATRIX PRODUCT

Math 21b, O. Knill

HOMEWORK: Section 2.4: 4,14,28,40,76,48*60*

MATRIX PRODUCT. If $B$ is a $p \times m$ matrix and $A$ is a $m \times n$ matrix, then $BA$ is defined as the $p \times n$ matrix with entries $(BA)_{ij} = \sum_{k=1}^{m} B_{ik} A_{kj}$.

EXAMPLE. If $B$ is a $3 \times 4$ matrix, and $A$ is a $4 \times 2$ matrix then $BA$ is a $3 \times 2$ matrix.

$$B = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 1 & 8 & 1 \\ 1 & 0 & 9 & 2 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad BA = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 1 & 8 & 1 \\ 1 & 0 & 9 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 15 & 13 \\ 14 & 11 \\ 10 & 5 \end{bmatrix}.$$

COMPOSING LINEAR TRANSFORMATIONS. Let us associate to the computer network (shown at the left) a linear transformation, then their composition $T \circ S : x \mapsto BA(x)$ is a linear transformation from $R^n$ to $R^n$.

EXAMPLE. Find the matrix which is a composition of a rotation around the $x$-axes by an angle $\pi/2$ followed by a rotation around the $z$-axes by an angle $\pi/2$.

SOLUTION. The first transformation has the property that $e_1 \mapsto e_1, e_2 \mapsto -e_2, e_3 \mapsto e_3, e_4 \mapsto -e_4$. If $A$ is the matrix belonging to the first transformation and $B$ the second, then $BA$ is the matrix to the composition.

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

EXAMPLE. A rotation dilation is the composition of a rotation by $\alpha = \arctan(b/a)$ and a dilation (scale) by $r = \sqrt{a^2 + b^2}$.

REMARK. Matrix multiplication is a generalization of usual multiplication of numbers or the dot product.

MATRIX ALGEBRA. Note that $AB \neq BA$ in general! Otherwise, the same rules apply as for numbers: $A(BC) = (AB)C$, $AA^{-1} = A^{-1}A = I_n$, $(AB)^{-1} = B^{-1}A^{-1}$, $A(B + C) = AB + AC$, $(B + C)A = BA + CA$ etc.

PARTITIONED MATRICES. The entries of matrices can themselves be matrices. If $B$ is a $m \times n$ matrix and $A$ is a $n \times p$ matrix, and assume the entries are $k \times k$ matrices, then $BA$ is a $m \times p$ matrix where each entry $(BA)_{ij} = \sum_{k=1}^{n} B_{ik} A_{kj}$ is a $k \times k$ matrix. Partitioning matrices can be useful to improve the speed of matrix multiplication (i.e. Strassen algorithm).

EXAMPLE. If $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$, where $A_{ij}$ are $k \times k$ matrices with the property that $A_{11}$ and $A_{22}$ are invertible, then $B = \begin{bmatrix} A_{11}^{-1} & -A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$ is the inverse of $A$.

APPLICATIONS. (The material which follows is for motivation purposes only, more applications appear in the homework).

NETWORKS. Let us associate to the computer network (shown at the left) a matrix

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$ To a worm in the first computer we associate a vector

$$x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \quad \text{The vector } Ax \text{ has a 1 at the places, where the worm could be in the next step.}$$

The vector $(AA)(x)$ tells, in how many ways the worm can go from the first computer to other hosts in 2 steps. In our case, it can go in three different ways back to the computer itself.

Matrices help to solve combinatorial problems (see movie "Good will Hunting"). For example, what does $(A^{1000})(x)$ tell about the worm infection of the network? What does it mean if $A^{1000}$ has no zero entries?

FRACTALS. Closely related to linear maps are affine maps $x \mapsto Ax + b$. They are compositions of a linear map with a translation. It is not a linear map if $B(0) \neq 0$. Affine maps can be disguised as linear maps in the following way: let $y = \begin{bmatrix} x \\ 1 \end{bmatrix}$ and define the $(n+1) \times (n+1)$ matrix $B = \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix}$. Then $By = \begin{bmatrix} Ax + b \\ 1 \end{bmatrix}$.

Fractals can be constructed by taking for example 3 affine maps $R, S, T$ which contract area. For a given object $Y_0$ define $Y_1 = R(Y_0) \cup S(Y_0) \cup T(Y_0)$ and recursively $Y_i = R(Y_{i-1}) \cup S(Y_{i-1}) \cup T(Y_{i-1})$. The above picture shows $Y_4$ after some iterations. In the limit, for example if $R(Y_0), S(Y_0)$ and $T(Y_0)$ are disjoint, the sets $Y_k$ converge to a fractal, an object with dimension strictly between 1 and 2.

CHAOs. Consider a map in the plane like $T : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 2x + 2 \sin(x) - y \\ y \end{bmatrix}$. We apply this map again and again and follow the points $(x_1, y_1) = T(x, y), (x_2, y_2) = T(T(x, y), y_2), \ldots$. One writes $T^n$ for the $n$-th iteration of the map and $(x_n, y_n)$ for the image of $(x, y)$ under the map $T^n$. The linear approximation of the map at a point $(x, y)$ is the matrix $DT(x, y) = \begin{bmatrix} 2 + 2 \cos(x) & -1 \\ 1 & 0 \end{bmatrix}$. The row vectors of $DT(x, y)$ are just the gradients of $f$ and $g$. $T$ is called chaotic at $(x, y)$, if the entries of $DT(T^n)(x, y)$ grow exponentially fast with $n$. By the chain rule, $DT(T^n)$ is the product of matrices $DT(x, y_i)$. For example, $T$ is chaotic at $(0, 0)$. If there is a positive probability to hit a chaotic point, then $T$ is called chaotic.
FALSE COLORS. Any color can be represented as a vector \((r, g, b)\), where \(r \in [0, 1]\) is the red, \(g \in [0, 1]\) is the green, and \(b \in [0, 1]\) is the blue component. Changing colors in a picture means applying a transformation on the cube. Let \(T : (r, g, b) \rightarrow (g, b, r)\) and \(S : (r, g, b) \rightarrow (r, g, 0)\). What is the composition of these two linear maps?

OPTICS. Matrices help to calculate the motion of light rays through lenses. A light ray \(y(s) = x + ms\) in the plane is described by a vector \((x, m)\). Following the light ray over a distance of length \(L\) corresponds to the map \((x, m) \rightarrow (x + mL, m)\). In the lens, the ray is bent depending on the height \(x\). The transformation in the lens is \((x, m) \rightarrow (x, m - kx)\), where \(k\) is the strength of the lens.

\[
\begin{bmatrix} x \\ m \end{bmatrix} \rightarrow A_L \begin{bmatrix} x \\ m \end{bmatrix} = \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ m \end{bmatrix}
\begin{bmatrix} x \\ m \end{bmatrix} \rightarrow B_k \begin{bmatrix} x \\ m \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -k & 1 \end{bmatrix} \begin{bmatrix} x \\ m \end{bmatrix}.
\]

Examples:
1) Eye looking far: \(A_B B_k\). 2) Eye looking at distance \(L\): \(A_B B_k A_L\).
3) Telescope: \(B_k A_L B_k\). (More about it in problem 80 in section 2.4).