Topics: Vector spaces, solving $Ax = b$, matrix transpose and inverse, orthogonal matrices, block matrices.

1. Recognizing vector spaces

A (real) vector space is a set V equipped with two operations:

- An operation $+$ (vector addition) which takes $v_1, v_2 \in V$ and outputs $v_1 + v_2 \in V$.
- An operation $\cdot$ (scalar multiplication) which takes $c \in \mathbb{R}$ and $v \in V$ and outputs $c \cdot v \in V$.

(NB: Part of these requirements is that the sum must actually lie in V, in which case we say that "V is closed under addition," and similarly V must be "closed under scalar multiplication.") These operations must also satisfy some axioms.[1] Here are two of them:

- Existence of a 'zero': there must exist $v_0 \in V$ such that $v_0 + v = v$ for all $v \in V$.
- Existence of additive inverses: for any $v \in V$, there must exist $v' \in V$ such that $v + v' = v_0$.

The 'zero' vector v_0 is usually just denoted 0, and the additive inverse of v is usually denoted $-v$.

Problem 1. In the vector space consisting of all functions on the real line, what is the zero vector?

Problem 2. Consider the set V consisting of pairs (a, b) of real numbers satisfying $a + 2b = 0$. How should one define the operations of 'vector addition' and '(real) scalar multiplication' on V? Does this make V into a vector space? What about when V is the set of pairs (a, b) of real numbers satisfying $a \geq 0$?

2. Solving $Ax = b$

The following four problems are equivalent:

- Basic math: find $x, y \in \mathbb{R}$ such that
$$\begin{aligned} x + 2y &= 5 \\ 4x + y &= 6 \end{aligned}$$
- Matrix form: find $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ such that
$$\begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$
- Column view: find $x, y \in \mathbb{R}$ such that
$$x \begin{pmatrix} 1 \\ 4 \end{pmatrix} + y \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$
- Row view: find $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ such that
  - Its dot product with $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is equal to 5.
  - Its dot product with $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$ is equal to 6.

One way to measure the complexity of an algorithm is to count the number of arithmetic operations (plus, subtract, multiply, divide) that go into it.

Problem 3. Consider a 2-variable system of equations
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_2 &= b_2 \end{aligned}$$
where x_1, x_2 are the unknowns. Solve it using back substitution, i.e. solve for x_2, then substitute to find x_1. How many arithmetic operations do you need? Can you guess how many operations are needed for an n-variable system of the same "upper triangular" shape?

[1] For a full list, see `https://en.wikipedia.org/wiki/Vector_space#Definition`.

3. Matrix transpose and inverse

Given an $m \times n$ matrix A, the transpose $A^\top$ is an $n \times m$ matrix whose rows are the columns of A. We have $(AB)^\top = B^\top A^\top$.

The matrix A is *orthogonal* if $A^\top A = I_{n\times n}$. Nonobvious fact: this implies that $m \geq n$, i.e. A is square or 'tall and skinny'. If the n columns of A are thought of as individual column vectors $v_1, \dots, v_n$, then A is orthogonal if and only if

$$v_i \cdot v_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

In this case, $v_1, \dots, v_n$ are called an *orthonormal* collection of vectors.

Problem 4. Find an orthogonal matrix Q whose rows do not form an orthonormal collection of vectors.

Problem 5. Fill in the question marks to make an orthogonal matrix:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & ? \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & ? \\ 0 & -\frac{1}{\sqrt{3}} & ? \end{pmatrix}$$

How many ways are there?

Given an $n \times n$ matrix A, an $n \times n$ matrix B satisfying $AB = I_{n\times n}$ is called an *inverse* of A. (Nonobvious fact: this condition implies $BA = I_{n\times n}$, and vice versa.) Any square matrix either has a unique inverse or no inverse at all. The inverse, if it exists, is denoted A^{-1}, and we say that A is *invertible*.

Nonobvious fact: a square matrix A is invertible if and only if $\det(A) \neq 0$.

Problem 6. Given *invertible* $n \times n$ matrices A and B, show that $(AB)^{-1}$ exists and equals $B^{-1}A^{-1}$.

Problem 7. Show that a square orthogonal matrix A is invertible, and satisfies $AA^\top = I_{n\times n}$ and $A^{-1} = A^\top$.

Problem 8. Show that the rows of a *square* orthogonal matrix form an orthonormal collection of vectors.

4. Block matrices

A *block decomposition* of an $m \times n$ matrix A is a way of writing A as a 'matrix of matrices':

$$A = \begin{pmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{n1} & \cdots & B_{nn} \end{pmatrix}$$

where each B_{ij} is a matrix. The B_{ij} don't need to all have the same size, but they need to be sized so as to fit together into an $m \times n$ grid.

Fact. Bock decompositions are compatible with matrix multiplication, for example:

If $A = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ and $C = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$,

then $AC = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} = \begin{pmatrix} B_{11}D_{11} + B_{12}D_{21} & B_{11}D_{12} + B_{12}D_{22} \\ B_{21}D_{11} + B_{22}D_{21} & B_{21}D_{12} + B_{22}D_{22} \end{pmatrix}$,

provided that the blocks are the right sizes for the matrix multiplications to be possible.

Problem 9. Does there exist some block decomposition of some matrix whose set of blocks is as follows: $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 & 4 \end{pmatrix}, \begin{pmatrix} 5 & 6 \end{pmatrix}$.

Problem 10. If $A = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ is a block decomposition, what is the corresponding block decomposition of $A^\top$?

5. Solutions

(1) The zero vector is the function which always outputs zero, i.e. $f(x) = 0$.

(2) Addition is $(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$. Scalar multiplication is $c(a, b) := (ca, cb)$. This does make V into a vector space, because $a_1 + 2b_1 = 0$ and $a_2 + 2b_2 = 0$ implies

$$(a_1 + a_2) + 2(b_1 + b_2) = 0$$

so V is closed under addition. Similarly, $a+2b = 0$ implies that $(ca)+2(cb) = 0$, so V is closed under scalar multiplication. The zero element is $(0, 0)$, and the additive inverse of (a, b) is $(-1)(a, b) = (-a, -b)$.

On the other hand, the set of pairs (a, b) with $a \geq 0$ is not a vector space. For example, it is not closed under scalar multiplication, because $(-1)(1, 0) = (-1, 0)$, which lies outside the indicated set.

(3) The second equation gives $x_2 = b_2/a_{21}$. This uses one operation (division). Next, we solve for x_1 as follows:

$$x_1 = \frac{b_1 - a_{12}x_2}{a_{11}}.$$

Since we already know x_2, we don't need to redo the division from before. So this involves a multiplication, a subtraction, and a division by a_{11}, which is three operations. Thus, four operations are required.

A similar n-variable upper-triangular system of equations would require n^2 operations to solve.

(4) In fact, any non-square orthogonal matrix works. This is because a non-square orthogonal matrix Q must be of size $m \times n$ with $m > n$ (by the 'nonobvious fact'), so $Q^\top$ is of size $n \times m$, so it cannot be orthogonal (again by the 'nonobvious fact'), so its columns are not an orthonormal collection. But the columns of $Q^\top$ are the rows of Q, so we conclude that the rows of Q are not an orthonormal collection.

On the other hand, see Problem 8, which says that if Q is square orthogonal, then the rows *are* an orthonormal collection.

(5) Label the unknown entries as follows:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & a \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & b \\ 0 & -\frac{1}{\sqrt{3}} & c \end{pmatrix}$$

If this matrix is to be orthogonal, then the dot product of the third column with either of the first two columns must be zero. This gives the equations $a + b = 0$ and $a - b - c = 0$ (upon clearing denominators). We conclude that $b = -a$ and $c = 2a$.

Furthermore, the dot product of the third column with itself must be one. This gives $6a^2 = 1$, so $a = \pm\frac{1}{\sqrt{6}}$. Either solution works, so there are two ways to make such an orthogonal matrix.

(6) We have

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AI_{n\times n}A^{-1} \\ &= AA^{-1} \\ &= I_{n\times n} \end{aligned}$$

where we have used the associativity of matrix multiplication. This shows that $B^{-1}A^{-1}$ is a one-sided inverse to AB, and by the 'nonobvious fact' we know that $B^{-1}A^{-1}$ is the unique two-sided inverse to AB, since AB is a square matrix.

(7) Since A is orthogonal, we have $A^\top A = I_{n\times n}$. Since A is square, we conclude that $A^\top$ is the inverse of A. By the 'nonobvious fact', the equation $A^\top A = I_{n\times n}$ also implies $AA^\top = I_{n\times n}$, i.e. a one-sided inverse is a two-sided inverse.

(8) From the previous problem, we know that a square orthogonal matrix A also satisfies $AA^\top = I_{n\times n}$. By viewing A as built out of row vectors, we see that this matrix multiplication involves taking the dot products of those row vectors, and so this equation says that the rows of A are an orthonormal collection.

(9) No. Although it is possible to arrange these little matrices into a grid, e.g.

$$\begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & 6 \end{pmatrix}$$

the definition of 'block decomposition' requires that the blocks B_{ij} themselves are arranged in a (smaller) grid. In other words, the lines separating the B_{ij} from each other must go all the way from the top to the bottom of the matrix, and all the way from the left to the right.

(10) We have

$$A^\top = \begin{pmatrix} B_{11}^\top & B_{21}^\top \\ B_{12}^\top & B_{22}^\top \end{pmatrix}.$$