Part I : Highlights of Linear Algebra

I.1 Multiplication Ax Using Columns of A

We hope you already know some linear algebra. It is a beautiful subject—more useful to more people than calculus (in our quiet opinion). But even old-style linear algebra courses miss basic and important facts. This first section of the book is about *matrix-vector* multiplication $A\boldsymbol{x}$.

We always use examples to make our point clear.

Example 1 Multiply A times $\boldsymbol{x}$ using the three rows of A and then using the two columns :

By rows $$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix} = \begin{array}{c} \text{inner products} \\ \text{of the rows} \\ \text{with } \boldsymbol{x} = (x_1, x_2) \end{array}$$

By columns $$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} = \begin{array}{c} \text{combination} \\ \text{of the columns} \\ \boldsymbol{a}_1 \text{ and } \boldsymbol{a}_2 \end{array}$$

You see that both ways give the same result. The first way (a row at a time) produces 3 inner products. Those are also known as "dot products" because of the dot notation :

$$\textbf{row} \cdot \textbf{column} = (\mathbf{2},\ \mathbf{3}) \cdot (\boldsymbol{x}_1,\ \boldsymbol{x}_2) = \mathbf{2}\boldsymbol{x}_1 + \mathbf{3}\boldsymbol{x}_2 \tag{1}$$

This is the way to find the three separate components of $A\boldsymbol{x}$. We use this for computing—but not for understanding. It is low level. Understanding is higher level, staying with vectors.

The vector approach sees $A\boldsymbol{x}$ as a "linear combination" of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. This is the fundamental operation of linear algebra ! A linear combination of vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ includes two steps :

(1) Multiply $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ by "scalars" x_1 and x_2

(2) Add vectors $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$.

Thus $A\boldsymbol{x}$ is a linear combination of the columns of A. This is fundamental.

This thinking leads us to the **column space** of A. The key idea is to take **all combinations** of the columns. All real numbers x_1 and x_2 are allowed—the space includes $A\boldsymbol{x}$ for all vectors $\boldsymbol{x}$. In this way we get infinitely many output vectors $A\boldsymbol{x}$. And we can see those outputs geometrically.

In our example, each $A\boldsymbol{x}$ is a vector in 3-dimensional space. That 3D space is called $\mathbf{R}^3$. (The $\mathbf{R}$ indicates real numbers. Vectors with three complex components lie in the space $\mathbf{C}^3$.) We stay with real vectors and we ask this key question :

All combinations $A\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$ produce what part of the full 3D space ?

Answer : Those vectors produce a **plane**. The plane contains the complete line in the direction of $\boldsymbol{a}_1 = (2, 2, 3)$, since every vector $x_1\boldsymbol{a}_1$ is included. The plane also includes the infinite line of all vectors $x_2\boldsymbol{a}_2$ in the direction of $\boldsymbol{a}_2$. And it includes the *sum* of any vector on one line plus any vector on the other line. **This addition fills out an infinite plane containing the two lines**. But it does not fill out the whole 3-dimensional space $\mathbf{R}^3$:

Definition **The combinations of the columns** fill out (**"span"**) the **column space** of A.

Here the column space is a plane. That plane includes the zero point $(0, 0, 0)$ which is produced when $x_1 = x_2 = 0$. The plane includes $(5, 6, 10) = \boldsymbol{a}_1 + \boldsymbol{a}_2$ and $(-1, -2, -4) = \boldsymbol{a}_1 - \boldsymbol{a}_2$. With probability 1 it does **not** include the random point **rand**$(3, 1)$! Which points are in the plane ?

$\boldsymbol{b} = (b_1, b_2, b_3)$ is in the column space of $\boldsymbol{A}$ **exactly when** $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a solution (x_1, x_2)

When you see that truth, you understand the column space $\mathbf{C}(A)$: The solution $\boldsymbol{x}$ shows how to express the right side $\boldsymbol{b}$ as a combination $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$ of the columns. For some $\boldsymbol{b}$ this is impossible.

Example 2 $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is not in $\mathbf{C}(A)$ because $A\boldsymbol{x} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is unsolvable.

The first two equations force $x_1 = \frac{1}{2}$ and $x_2 = 0$. Then equation 3 fails : $3\left(\frac{1}{2}\right) + 7(0) = 1.5$ (**not 1**).

This means that $\boldsymbol{b} = (1, 1, 1)$ is not in the column space—the plane of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.

Example 3 What are the column spaces of $A_2 = \begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix}$ and $A_3 = \begin{bmatrix} 2 & 3 & 1 \\ 2 & 4 & 1 \\ 3 & 7 & 1 \end{bmatrix}$?

Solution. The column space of A_2 is the same plane as before. The new column $(5, 6, 10)$ is the sum of column 1 + column 2. So $\boldsymbol{a}_3$ = column 3 is already in the plane and adds nothing new. By including this "*dependent*" column we don't go beyond the original plane $\mathbf{C}(A)$.

The column space of A_3 is the whole 3D space $\mathbf{R}^3$. Example 2 showed us that the new third column $(1, 1, 1)$ is not in the plane $\mathbf{C}(A)$. Our column space $\mathbf{C}(A_3)$ has grown bigger. But there is nowhere to stop between a plane and the full 3D space. Visualize the $x - y$ plane and a third vector (x_3, y_3, z_3) out of the plane (meaning that $z_3 \neq 0$). They combine to give **every vector in $\mathbf{R}^3$**.

Here is a total and exclusive list of all possible column spaces inside $\mathbf{R}^3$. Dimensions $0, 1, 2, 3$:

Subspaces of $\mathbf{R}^3$
- The **zero vector** $(0, 0, 0)$ by itself
- A **line** of all vectors $x_1\boldsymbol{a}_1$
- A **plane** of all vectors $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$
- The **whole $\mathbf{R}^3$** with all vectors $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3$

In that list we need the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ to be "**independent**". The only combination that gives the zero vector is $0\boldsymbol{a}_1 + 0\boldsymbol{a}_2 + 0\boldsymbol{a}_3$. So $\boldsymbol{a}_1$ by itself gives a line, $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ give a plane, $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ give every vector $\boldsymbol{b}$ in $\mathbf{R}^3$. The zero vector is in every subspace ! In linear algebra language :

- Three independent columns in $\mathbf{R}^3$ produce an **invertible** 3×3 matrix : $AA^{-1} = A^{-1}A = I$.
- $A\boldsymbol{x} = \mathbf{0}$ requires $\boldsymbol{x} = (0, 0, 0)$. $A\boldsymbol{x} = \boldsymbol{b}$ has exactly one solution $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ for every $\boldsymbol{b}$.

You see the picture for the columns of an n by n invertible matrix. Their combinations fill all of $\mathbf{R}^n$. Then the n rows will also be independent. We needed those ideas and that language to go further.

Problem Set I.1

1 Give an example where a combination of three nonzero vectors in $\mathbf{R}^4$ is the zero vector. Then write your example in the form $A\boldsymbol{x} = \mathbf{0}$. What are the shapes of A and $\boldsymbol{x}$ and $\mathbf{0}$?

2 Suppose a combination of the columns of A equals a different combination of those columns. Write that statement as $A\boldsymbol{x} = A\boldsymbol{y}$. Find a combination of the columns of A that equals the zero vector (in matrix language, find a solution to $A\boldsymbol{z} = \mathbf{0}$). Then find a second solution $\boldsymbol{z}_2$.

3 (Practice with subscripts) The vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ are in m-dimensional space $\mathbf{R}^m$, and a combination $c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n$ is the zero vector. That statement is at the vector level.

(1) Write that statement at the matrix level. Use the matrix A with the $\boldsymbol{a}$'s in its columns and use the column vector $\boldsymbol{c} = (c_1, \ldots, c_n)$.

(2) Write that statement at the scalar level using subscripts and sigma notation to add up numbers. The column vector $\boldsymbol{a}_j$ has components $a_{1j}, a_{2j}, \ldots, a_{mj}$.

4 Suppose A is the 3 by 3 matrix **ones**$(3, 3)$ of all ones. Find two independent vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ that solve $A\boldsymbol{x} = \mathbf{0}$ and $A\boldsymbol{y} = \mathbf{0}$. Write that first equation $A\boldsymbol{x} = \mathbf{0}$ (with numbers) as a combination of the columns of A. Why don't I ask for a third independent vector with $A\boldsymbol{z} = \mathbf{0}$?

5 The linear combinations of $\boldsymbol{v} = (1, 1, 0)$ and $\boldsymbol{w} = (0, 1, 1)$ fill a plane in $\mathbf{R}^3$.

(a) Find a vector $\boldsymbol{z}$ that is perpendicular to $\boldsymbol{v}$ and $\boldsymbol{w}$. Then $\boldsymbol{z}$ is perpendicular to every vector $c\boldsymbol{v} + d\boldsymbol{w}$ on the plane : $(c\boldsymbol{v} + d\boldsymbol{w})^{\mathrm{T}} z = c\boldsymbol{v}^{\mathrm{T}}\boldsymbol{z} + d\boldsymbol{w}^{\mathrm{T}} z = 0 + 0$.

(b) Find a vector $\boldsymbol{u}$ that is not on the plane. Check that $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{z} \neq 0$.

6 In the xy plane mark all nine of these linear combinations:

$$c \begin{bmatrix} 2 \\ 1 \end{bmatrix} + d \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{with} \quad c = 0, 1, 2 \quad \text{and} \quad d = 0, 1, 2.$$

7 If three corners of a parallelogram are $(1, 1)$, $(4, 2)$, and $(1, 3)$, what are all three of the possible fourth corners? Draw two of them.

8 Draw two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ coming out from the center point $(0, 0)$.

(a) Mark the points $\frac{1}{2}\boldsymbol{v} + \frac{1}{2}\boldsymbol{w}$ and $\frac{1}{4}\boldsymbol{v} + \frac{1}{4}\boldsymbol{w}$.

(b) Draw a line containing all the points $c\boldsymbol{v} + (1 - c)\,\boldsymbol{w}$ (all c).

(c) Draw the "cone" of all combinations $c\boldsymbol{v} + d\boldsymbol{w}$ with $c \geq 0$ and $d \geq 0$.

9 How many corners does a cube have in 4 dimensions ? How many 3D faces ? How many edges ? An edge goes between two adjacent corners.

10 Describe the column space of $A = [\boldsymbol{v} \;\; \boldsymbol{w} \;\; \boldsymbol{v} + 2\boldsymbol{w}]$. Describe the nullspace of A : all vectors $\boldsymbol{x} = (x_1, x_2, x_3)$ that solve $A\boldsymbol{x} = \mathbf{0}$. Add the "dimensions" of that plane and that line :

dimension of column space $+$ dimension of nullspace $=$ number of columns

11 Suppose the column space of an m by n matrix is all of $\mathbf{R}^3$. What can you say about m ? What can you say about n ?

I.2 Matrix-Matrix Multiplication AB

Inner products (*rows times columns*) produce each of the numbers in $AB = C$:

$$\begin{matrix}\textbf{row 2 of } \boldsymbol{A} \\ \textbf{column 3 of } \boldsymbol{B} \\ \textbf{give } \boldsymbol{c_{23}} \textbf{ in } \boldsymbol{C}\end{matrix} \qquad \begin{bmatrix} \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot & b_{13} \\ \cdot & \cdot & b_{23} \\ \cdot & \cdot & b_{33} \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \boldsymbol{c_{23}} \\ \cdot & \cdot & \cdot \end{bmatrix} \tag{1}$$

That dot product $c_{23} =$ (row 2 of A) • (column 3 of B) is a sum:

$$c_{23} = a_{\mathbf{21}}\, b_{\mathbf{13}} + a_{\mathbf{22}}\, b_{\mathbf{23}} + a_{\mathbf{23}}\, b_{\mathbf{33}} = \sum_{k=1}^{3} a_{2k}\, b_{k3} \quad \text{and} \quad c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}. \tag{2}$$

This is how we usually compute each number in $AB = C$.

There is another way to multiply AB: **columns of $\boldsymbol{A}$ times rows of $\boldsymbol{B}$**. We need to see this! I start with numbers to make two key points: one column $\boldsymbol{u}$ times one row $\boldsymbol{v}^{\mathrm{T}}$ produces a *matrix*. Concentrate first on that piece of AB. This matrix $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ is especially simple:

$$\begin{matrix}\textbf{Outer} \\ \textbf{product}\end{matrix} \qquad \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 8 & 12 \\ 6 & 8 & 12 \\ 3 & 4 & 6 \end{bmatrix} = \begin{matrix}\textbf{``rank one} \\ \textbf{matrix''}\end{matrix}$$

An m by 1 matrix (a column $\boldsymbol{u}$) times a 1 by p matrix (a row $\boldsymbol{v}^{\mathrm{T}}$) gives an m by p matrix. Notice what is special about the rank one matrix $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$:

$$\text{All columns of } \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} \text{ are multiples of } \boldsymbol{u} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}. \text{ All rows are multiples of } \boldsymbol{v}^{\mathrm{T}} = \begin{bmatrix} 3 & 4 & 6 \end{bmatrix}.$$

The column space of $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ is one-dimensional: *the line in the direction of* $\boldsymbol{u}$. The dimension of the column space (the number of independent columns) is the **rank of the matrix**—a key number. **All nonzero matrices $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ have rank one**. They are the perfect building blocks for every matrix.

Notice also: **The row space of $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ is the line through $\boldsymbol{v}$**. By definition, the row space of any matrix A is the column space $\mathbf{C}(A^{\mathrm{T}})$ of its transpose A^{T}. That way we stay with column vectors. In the example, we transpose $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ (**exchange rows with columns**) to get the matrix $\boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}$:

$$(\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{\mathrm{T}} = \begin{bmatrix} \mathbf{6} & 8 & 12 \\ \mathbf{6} & 8 & 12 \\ \mathbf{3} & 4 & 6 \end{bmatrix}^{\mathbf{T}} = \begin{bmatrix} \mathbf{6} & \mathbf{6} & \mathbf{3} \\ 8 & 8 & 4 \\ 12 & 12 & 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \end{bmatrix} = \boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}.$$

We are seeing the clearest possible example of the first great theorem in linear algebra:

Row rank = Column rank	**$\boldsymbol{r}$ independent columns ⇔ $\boldsymbol{r}$ independent rows**

A nonzero matrix $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ has one independent column and one independent row! All columns are multiples of $\boldsymbol{u}$ and all rows are multiples of $\boldsymbol{v}^{\mathrm{T}}$. The rank is $r = 1$ for this matrix.

AB = Sum of Rank One Matrices

We turn to the full product AB, using columns of A times rows of B. Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ be the columns of A. Let $\boldsymbol{b}_1^*, \boldsymbol{b}_2^*, \ldots, \boldsymbol{b}_n^*$ be the rows of B. Notice the same number n (or we couldn't multiply A times B). **Then the product $\boldsymbol{AB}$ is the sum of columns $\boldsymbol{a}_k$ times rows $\boldsymbol{b}_k^*$:**

Column-row multiplication of matrices

$$AB = \begin{bmatrix} | & & | \\ \boldsymbol{a}_1 & \ldots & \boldsymbol{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} — & \boldsymbol{b}_1^* & — \\ & \vdots & \\ — & \boldsymbol{b}_n^* & — \end{bmatrix} = \boldsymbol{a}_1\,\boldsymbol{b}_1^* + \boldsymbol{a}_2\,\boldsymbol{b}_2^* + \cdots + \boldsymbol{a}_n\,\boldsymbol{b}_n^*. \tag{3}$$

Here is a 2 by 2 example to show the $n = 2$ pieces (column times row) and their sum AB:

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{3} \end{bmatrix} \begin{bmatrix} \mathbf{2} & \mathbf{4} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{5} \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 12 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 17 \end{bmatrix}. \tag{4}$$

Can you count the multiplications of number times number? Four multiplications to get $2, 4, 6, 12$. Four more to get $0, 0, 0, 5$. A total of $2^3 = 8$ multiplications. Always there are $\boldsymbol{n^3}$ multiplications when A and B are n by n. And $\boldsymbol{mnp}$ multiplications when $AB = (m \text{ by } n)$ times $(n \text{ by } p)$: n rank one matrices, each of those matrices is m by p.

The count is the same for the usual inner product way! Row of A times column of B needs n multiplications. We do this for every number in AB: mp dot products when AB is m by p. The total count is again $\boldsymbol{mnp}$ for $(m$ by $n)$ times $(n$ by $p)$.

$\boldsymbol{mp}$ inner products, $\boldsymbol{n}$ multiplications each OR $\boldsymbol{n}$ outer products, $\boldsymbol{mp}$ multiplications each

When you look closely, they are exactly the same multiplications $a_{ik}\,b_{kj}$ in different orders. Here is the algebra proof that c_{ij} is the same by outer products in (3) as by inner products in (2):

The i, j entry of $\boldsymbol{a}_k\boldsymbol{b}_k^*$ is $\boldsymbol{a_{ik}b_{kj}}$. Add for $k = 1$ to n. Then $c_{ij} = \sum_{k=1}^{n} a_{ik}\,b_{kj} =$ **row $\boldsymbol{i}$ · column $\boldsymbol{j}$**.

Insight from Column times Row

Why is the outer product approach essential in data science? The short answer is: *We are looking for the important part of a matrix C.* We don't usually want the biggest number in C (though that could be important). What we want more is the largest piece of C. **And those pieces are rank one matrices $\boldsymbol{uv}^{\mathrm{T}}$.** A dominant theme in applied linear algebra is:

Factor $\boldsymbol{C}$ into $\boldsymbol{AB}$ and look at the pieces $\boldsymbol{a_k b_k^*}$ of $\boldsymbol{AB = C}$.

Factoring C into AB is the reverse of multiplying $AB = C$. Factoring takes longer, especially if the pieces involve *eigenvalues* or *singular values*. But those numbers have inside information about the matrix C. That information is not visible until you factor.

Here are five important factorizations, written with the standard choice of letters (usually A, not C) for the original product matrix and then for its factors. This book will explain all five.

$$\boldsymbol{A = LU} \qquad \boldsymbol{A = QR} \qquad \boldsymbol{S = Q\Lambda Q^{\mathrm{T}}} \qquad \boldsymbol{A = X\Lambda X^{-1}} \qquad \boldsymbol{A = U\Sigma V^{\mathrm{T}}}$$

At this point we simply list key words and properties for each of these factorizations.

1 $A = LU$ comes from **elimination**. Combinations of rows take A to U and U back to A
L is lower triangular and U is upper triangular as in equation (4)

2 $A = QR$ comes from **orthogonalizing** the columns $\boldsymbol{a}_1$ to $\boldsymbol{a}_n$ as in "Gram-Schmidt"
Q has orthonormal columns ($Q^{\mathrm{T}}Q = I$) and R is upper triangular

3 $S = Q\Lambda Q^{\mathrm{T}}$ comes from the **eigenvalues** $\lambda_1, \ldots, \lambda_n$ of a symmetric matrix $S = S^{\mathrm{T}}$
Eigenvalues on the diagonal of Λ and **orthonormal eigenvectors** in the columns of Q

4 $A = X\Lambda X^{-1}$ is **diagonalization** when A is n by n with n independent eigenvectors
Eigenvalues of A are on the diagonal of Λ. Eigenvectors of A are in the columns of X

5 $A = U\Sigma V^{\mathrm{T}}$ is the **Singular Value Decomposition** of any matrix A (square or not)
The **singular values** $\sigma_1, \ldots, \sigma_r$ are in Σ. The orthonormal **singular vectors** are in U and V

Let me pick out a favorite (number **3**) to illustrate the idea. This special factorization $Q\Lambda Q^{\mathrm{T}}$ starts with a symmetric matrix S. That matrix has orthogonal unit eigenvectors $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$. Those eigenvectors go into the columns of Q. S and Q are the kings and queens of linear algebra:

Symmetric matrix S $\quad \boldsymbol{S^{\mathrm{T}} = S}$ $\quad$ when all $\boldsymbol{s_{ij} = s_{ji}}$

Orthogonal matrix Q $\quad \boldsymbol{Q^{\mathrm{T}} = Q^{-1}}$ $\quad$ when all $\boldsymbol{q_i \cdot q_j} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$

The diagonal matrix Λ contains real eigenvalues λ_1 to λ_n. Every real symmetric matrix S has n orthonormal eigenvectors $\boldsymbol{q}_1$ to $\boldsymbol{q}_n$. When multiplied by S, they keep the same direction:

$$\textbf{Eigenvector } \boldsymbol{q} \textbf{ and eigenvalue } \boldsymbol{\lambda} \qquad S\boldsymbol{q} = \lambda\boldsymbol{q} \tag{5}$$

Finding λ and $\boldsymbol{q}$ is not easy for a big matrix. But n pairs always exist when S is symmetric. Our purpose here is to see how $SQ = Q\Lambda$ comes column by column from $S\boldsymbol{q} = \lambda\boldsymbol{q}$:

$$\boldsymbol{SQ} = S\begin{bmatrix} \boldsymbol{q}_1 & \ldots & \boldsymbol{q}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{q}_1 & \ldots & \lambda_n\boldsymbol{q}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}_1 & \ldots & \boldsymbol{q}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \boldsymbol{Q\Lambda} \tag{6}$$

Multiply $SQ = Q\Lambda$ by $Q^{-1} = Q^{\mathrm{T}}$ to get $\boldsymbol{S = Q\Lambda Q^{\mathrm{T}}}$ = a symmetric matrix. Each eigenvalue λ_k and each eigenvector $\boldsymbol{q}_k$ contribute a rank one piece $\lambda_k\boldsymbol{q}_k\boldsymbol{q}_k^{\mathrm{T}}$ to S.

$$\textbf{Always symmetric} \qquad (Q\Lambda Q^{\mathrm{T}})^{\mathrm{T}} = Q^{\mathrm{TT}}\Lambda^{\mathrm{T}}Q^{\mathrm{T}} = Q\Lambda Q^{\mathrm{T}} \quad (\text{diagonal } \Lambda^{\mathrm{T}} = \Lambda) \tag{7}$$

$$\textbf{Rank one pieces} \qquad S = (Q\Lambda)Q^{\mathrm{T}} = (\lambda_1\boldsymbol{q}_1)\boldsymbol{q}_1^{\mathrm{T}} + (\lambda_2\boldsymbol{q}_2)\boldsymbol{q}_2^{\mathrm{T}} + \cdots + (\lambda_n\boldsymbol{q}_n)\boldsymbol{q}_n^{\mathrm{T}} \tag{8}$$

Please notice that the columns of $Q\Lambda$ are $\lambda_1\boldsymbol{q}_1$ to $\lambda_n\boldsymbol{q}_n$. When you multiply a matrix on the right by the diagonal matrix Λ, you multiply its *columns* by the λ's.