Essential Mathematics for the Political and Social Research

JEFF GILL
Cambridge University Press

Lecture Slides, Chapter 3: Linear Algebra: Vectors, Matrices, and Operations

Chapter 3 Objectives

- When one reads high visibility journals in the social sciences, matrix algebra is ubiquitous.

- Matrix Algebra is useful because it makes our life easier since it lets us express extensive models in quite readable notation.

- Consider the following linear statistical model specification [from real work, Powers and Cox (1997)], relating political blame to various demographic and regional political variables:

$$\begin{aligned}\text{for } i = 1 \text{ to } n,\ (BLAMEFIRST)Y_i = &\\ \beta_0 + \beta_1 CHANGELIV + \beta_2 BLAMECOMM + \beta_3 INCOME &\\ + \beta_4 FARMER + \beta_5 OWNER + \beta_6 BLUESTATE &\\ + \beta_7 WHITESTATE + \beta_8 FORMMCOMM + \beta_9 AGE &\\ + \beta_{10} SQAGE + \beta_{11} SEX + \beta_{12} SIZEPLACE &\\ + \beta_{13} EDUC + \beta_{14} FINHS + \beta_{15} ED * HS &\\ + \beta_{16} RELIG + \beta_{17} NATION + E_i &\end{aligned}$$

- In matrix algebra form this is simply $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{E}$.

- Because we can notate large groups of numbers in an easy-to-read structural form, we can concentrate more on the theoretically interesting properties of the analysis.

Working with Vectors

- A vector is just a serial listing of numbers where the order matters.
- A vector of the first four positive integers in a single vector:

$$\text{a row vector: } \mathbf{v} = [1, 2, 3, 4], \qquad \text{or a column vector: } \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix},$$

  where $\mathbf{v}$ is the name for this new object.
- Order matters in the sense that the two vectors above are different, for instance, from

$$\mathbf{v}^* = [4, 3, 2, 1], \qquad \mathbf{v}^* = \begin{bmatrix} 4 \\ 2 \\ 3 \\ 1 \end{bmatrix}.$$

- It is a convention that vectors are designated in bold type and individual values, *scalars*, are designated in regular type.
- Thus $\mathbf{v}$ is a vector with elements v_1, v_2, v_3, v_4, and v would be some *other* scalar quantity.

A Note on the "Direction" of Vectors

- It does not matter whether we consider a vector to be of column or row form, unless we are using it in equations.
- Some disciplines (notably economics) tend to default to the column form:

$$\begin{bmatrix} 4 \\ 2 \\ 3 \\ 1 \end{bmatrix}.$$

- In the row form, it is equally common to see spacing used instead of commas as delimiters: $[1\ 2\ 3\ 4]$.

Arithmatic of Vectors

The following examples use the vectors $\mathbf{u} = [3, 3, 3, 3]$ and $\mathbf{v} = [1, 2, 3, 4]$.

- **Vector Addition Calculation**.

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, u_3 + v_3, u_4 + v_4] = [4, 5, 6, 7].$$

- **Vector Subtraction Calculation**.

$$\mathbf{u} - \mathbf{v} = [u_1 - v_1, u_2 - v_2, u_3 - v_3, u_4 - v_4] = [2, 1, 0, -1].$$

- **Scalar Multiplication Calculation**.

$$3 \times \mathbf{v} = [3 \times v_1, 3 \times v_2, 3 \times v_3, 3 \times v_4] = [3, 6, 9, 12].$$

- **Scalar Division Calculation**.

$$\mathbf{v} \div 3 = [v_1/3, v_2/3, v_3/3, v_4/3] = \left[\frac{1}{3}, \frac{2}{3}, 1, \frac{4}{3}\right].$$

Conditions for Arithmetic Operations on Vectors

- The key issue with addition or subtraction between two vectors is that the operation is applied only to the corresponding vector elements as pairs: the first vector elements together, the second vector elements together, and so on.
- So *the vectors have to be exactly the same size* (same number of elements).
- This is called *conformable* in the sense that the first vector must be of a size that conforms with the second vector.
- Otherwise they are called *nonconformable*.
- In the previous examples both $\mathbf{u}$ and $\mathbf{v}$ are 1×4 (row) vectors (alternatively called length $k = 4$ vectors), meaning that they have one row and four columns.
- Sometimes size is denoted beneath the vectors:

$$\underset{1\times 4}{\mathbf{u}} + \underset{1\times 4}{\mathbf{v}}.$$

Special Vectors

- $\mathbf{1}$: a vector of all 1's.
- $\mathbf{0}$: a vector of all 0's.
- There are a much larger number of "special" matrices that have similarly important characteristics.

Vector Calculation Properties

- For vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$, which are identically sized, and the scalars s and t. The following intuitive properties hold.

Elementary Formal Properties of Vector Algebra

$\rightarrowtail$ Commutative Property	$\mathbf{u} + \mathbf{v} = (\mathbf{v} + \mathbf{u})$
$\rightarrowtail$ Additive Associative Property	$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
$\rightarrowtail$ Vector Distributive Property	$s(\mathbf{u} + \mathbf{v}) = s\mathbf{u} + s\mathbf{v}$
$\rightarrowtail$ Scalar Distributive Property	$(s + t)\mathbf{u} = s\mathbf{u} + t\mathbf{u}$
$\rightarrowtail$ Zero Property	$\mathbf{u} + \mathbf{0} = \mathbf{u} \Longleftrightarrow \mathbf{u} - \mathbf{u} = \mathbf{0}$
$\rightarrowtail$ Zero Multiplicative Property	$\mathbf{0}\mathbf{u} = \mathbf{0}$
$\rightarrowtail$ Unit Rule	$\mathbf{1}\mathbf{u} = \mathbf{u}$

Vector Calculation Examples

- Define $s = 3$, $t = 1$, $\mathbf{u} = [2, 4, 8]$, and $\mathbf{v} = [9, 7, 5]$.
- Then:

$(s+t)(\mathbf{v}+\mathbf{u})$	$s\mathbf{v}+t\mathbf{v}+s\mathbf{u}+t\mathbf{u}$
$(3+1)([9,7,5]+[2,4,8])$	$3[9,7,5]+1[9,7,5]+3[2,4,8]+1[2,4,8]$
$4[11,11,13]$	$[27,21,15]+[9,7,5]+[6,12,24]+[2,4,8]$
$[44,44,52]$	$[44,44,52]$

Vector Inner Product

- The *inner product*, also called the *dot product*, of two vectors, results in a scalar (and so it is also called the *scalar product*).
- The inner product of two conformable vectors of arbitrary length k is the sum of the item-by-item products:

$$\mathbf{u} \cdot \mathbf{v} = [u_1 v_1 + u_2 v_2 + \cdots u_k v_k] = \sum_{i=1}^{k} u_i v_i.$$

- Note the use of the summation notation.
- **Simple Inner Product Calculation**. A numerical example of an inner product multiplication is given by

$$\mathbf{u} \cdot \mathbf{v} = [3, 3, 3] \cdot [1, 2, 3] = [3 \cdot 1 + 3 \cdot 2 + 3 \cdot 3] = 18.$$

Orthogonal Vectors

- When the inner product of two vectors is zero, we say that the vectors are *orthogonal*.
- This means they are at a right angle to each other
- The notation for the orthogonality of two vectors is $\mathbf{u} \perp \mathbf{v}$ iff $\mathbf{u} \cdot \mathbf{v} = 0$.
- As an example of orthogonality, consider $\mathbf{u} = [1, 2, -3]$, and $\mathbf{v} = [1, 1, 1]$.

Formal Properties for Inner Products

Inner Product Formal Properties of Vector Algebra

- $\rightarrowtail$ Commutative Property $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $\rightarrowtail$ Associative Property $s(\mathbf{u} \cdot \mathbf{v}) = (s\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (s\mathbf{v})$
- $\rightarrowtail$ Distributive Property $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- $\rightarrowtail$ Zero Property $\mathbf{u} \cdot \mathbf{0} = 0$
- $\rightarrowtail$ Unit Rule $\mathbf{1u} = \mathbf{u}$
- $\rightarrowtail$ Unit Rule $\mathbf{1u} = \sum_{i=1}^{k} \mathbf{u}_i$, for $\mathbf{u}$ of length k

Examples of Inner Product Calculations

- **Vector Inner Product Calculations**. This example demonstrates the first three properties above. Define $s = 5$, $\mathbf{u} = [2, 3, 1]$, $\mathbf{v} = [4, 4, 4]$, and $\mathbf{w} = [-1, 3, -4]$. Then:

$s(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w}$	$s\mathbf{v} \cdot \mathbf{w} + s\mathbf{u} \cdot \mathbf{w}$
$5([2, 3, 1] + [4, 4, 4]) \cdot [-1, 3, -4]$	$5[4, 4, 4] \cdot [-1, 3, -4]$
	$+5[2, 3, 1] \cdot [-1, 3, -4]$
$5([6, 7, 5]) \cdot [-1, 3, -4]$	$[20, 20, 20] \cdot [-1, 3, -4]$
	$+[10, 15, 5] \cdot [-1, 3, -4]$
$[30, 35, 25] \cdot [-1, 3, -4]$	$-40 + 15$
-25	-25

Vector Cross Product

- The *cross product* of two vectors, sometimes called the *outer product*, although this term is better reserved for a slightly different operation.
- This is slightly more involved than the inner product, in both calculation and interpretation because the result is not a scalar.
- The cross product of two conformable vectors of length $k = 3$ is

$$\mathbf{u} \times \mathbf{v} = \left[u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1 \right],$$

- This means that the first element is a difference equation that leaves out the first elements of the original two vectors, and the second and third elements proceed accordingly.
- In the more general sense, we perform a series of "leave one out" operations that is more extensive than above because the suboperations are themselves cross products.

Vector Cross Product Illustration

- First the **u** and **v** vectors are stacked on top of each other in the upper part of the illustration.
- The process of calculating the first vector value of the cross product, which we will call w_1, is done by "crossing" the elements in the solid box: u_2v_3 indicated by the first arrow and u_3v_2 indicated by the second arrow.
- The result for w_1 as a difference between these two individual components.
- This is actually the *determinant* of the 2×2 submatrix, which is an important principle considered later.

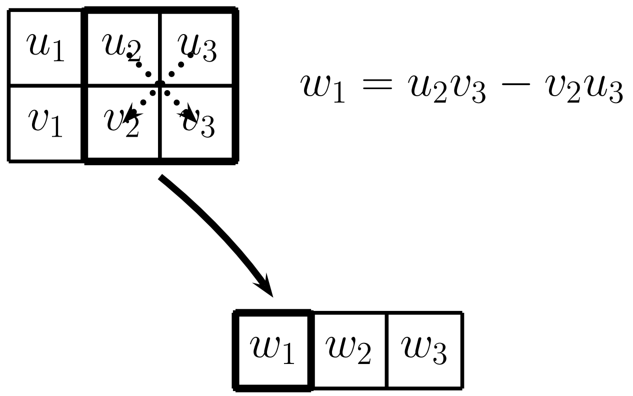
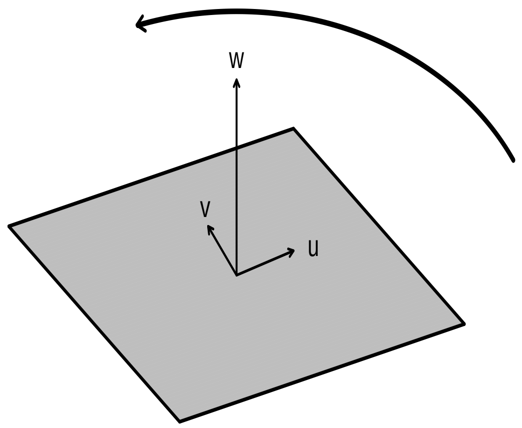


Figure 1: The Right-Hand Rule Illustrated



The Right-Hand Rule

- The resulting vector from a cross product is orthogonal to both of the original vectors in the direction of the so-called "right-hand rule."
- If you hold your hand as you would when hitchhiking, the curled fingers make up the original vectors and the thumb indicates the direction of the orthogonal vector that results from a cross product.
- In the figure you can imagine your right hand resting on the plane with the fingers curling to the left (⥀) and the thumb facing upward.

Properties of Cross Products

- For vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$, the cross product properties are given by:

Cross Product Formal Properties of Vector Algebra

$\rightarrowtail$ Commutative Property $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$

$\rightarrowtail$ Associative Property $s(\mathbf{u} \times \mathbf{v}) = (s\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (s\mathbf{v})$

$\rightarrowtail$ Distributive Property $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$

$\rightarrowtail$ Zero Property $\mathbf{u} \times \mathbf{0} = \mathbf{0} \times \mathbf{u} = \mathbf{0}$

$\rightarrowtail$ Self-Orthogonality $\mathbf{u} \times \mathbf{u} = \mathbf{0}$

Cross Product Example

- Returning to the simple numerical example from before, we now calculate the cross product instead of the inner product:

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= [3,3,3] \times [1,2,3] \\ &= [(3)(3)-(3)(2),(3)(1)-(3)(3),(3)(2)-(3)(1)] = [3,-6,3]. \end{aligned}$$

- We can then check the orthogonality as well:

$$[3,3,3] \cdot [3,-6,3] = 0 \qquad [1,2,3] \cdot [3,-6,3] = 0.$$

Back to Row Versus Column Vectors

- Vector multiplication should be done in a conformable manner with regard to multiplication when a row vector multiplies a column vector such that their adjacent "sizes" match: a $(1 \times k)$ vector multiplying a $(k \times 1)$ vector for k elements in each.
- This operation is now an inner product:

$$\underset{1\times k}{[v_1, v_2, \dots, v_k]} \times \underset{k\times 1}{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}}.$$

- This adjacency above comes from the k that denotes the columns of $\mathbf{v}$ and the k that denotes the rows of $\mathbf{u}$ and manner by which they are next to each other.

Back to Row Versus Column Vectors

- An outer product would be implied by this type of adjacency:

$$\underset{k\times 1}{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}} \times \underset{1\times k}{[v_1, v_2, \ldots, v_k]},$$

  where the 1's are next to each other.

- The cross product of two vectors is a vector, and the outer product of two conformable vectors is a matrix: a rectangular grouping of numbers that generalizes vectors.

Transpose

- To be completely explicit about these operations we can also use the *vector transpose*, which converts a row vector to a column vector, or vice versa, using the apostrophe notation:

$$\underset{k\times 1}{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}}' = \underset{1\times k}{[u_1, u_2, \ldots, u_k]}, \qquad \underset{1\times k}{[u_1, u_2, \ldots, u_k]}' = \underset{k\times 1}{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}}.$$

- This is just a notational device: it does not change the numerical value of the vector in any way.
- However, it is really important with matrix operations.
- Note that the order of multiplication now matters with the outer product.

Examples of Outer Product Calculations

- Once again using the simple numerical forms, we now calculate the outer product instead of the cross product:

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [3, 3, 3] = \begin{bmatrix} 3 & 3 & 3 \\ 6 & 6 & 6 \\ 9 & 9 & 9 \end{bmatrix}.$$

  And to show that order matters, consider:

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} [1, 2, 3] = \begin{bmatrix} 3 & 6 & 9 \\ 3 & 6 & 9 \\ 3 & 6 & 9 \end{bmatrix}.$$

Vector Norms

- Measuring the "length" of vectors can be done in different ways.
- For example , $(5,5,5)$ should be considered longer than $(1,1,1)$, but it is not clear whether $(4,4,4)$ is longer than $(3,-6,3)$.
- The standard version of the *vector norm* for an n-length vector is given by

$$\|\mathbf{v}\| = (v_1^2 + v_2^2 + \cdots + v_n^2)^{\frac{1}{2}} = (\mathbf{v}' \cdot \mathbf{v})^{\frac{1}{2}}.$$

- The vector norm can be thought of as the distance of the vector from the origin.
- Using the formula for $\|\mathbf{v}\|$ we can now calculate the vector norm for $(4,4,4)$ and $(3,-6,3)$:

$$\|(4,4,4)\| = \sqrt{4^2 + 4^2 + 4^2} = 6.928203$$

$$\|(3,-6,3)\| = \sqrt{3^2 + (-6)^2 + 3^2} = 7.348469.$$

Properties of Vector Norms

Properties of the Standard Vector Norm

$\rightarrowtail$ Vector Norm $||\mathbf{u}||^2 = \mathbf{u} \cdot \mathbf{u}$

$\rightarrowtail$ Difference Norm $||\mathbf{u} - \mathbf{v}||^2 = ||\mathbf{u}||^2 - 2(\mathbf{u} \cdot \mathbf{v}) + ||\mathbf{v}||^2$

$\rightarrowtail$ Multiplication Norm $||\mathbf{u} \times \mathbf{v}||^2 = ||\mathbf{u}||^2||\mathbf{v}||^2 - (\mathbf{u} \cdot \mathbf{v})^2$

Norm Calculation Examples

- Suppose $\mathbf{u} = [-10, 5]$ and $\mathbf{v} = [3, 3]$. **Difference Norm Calculations**:

$\lvert\lvert\mathbf{u} - \mathbf{v}\rvert\rvert^2$	$\lvert\lvert\mathbf{u}\rvert\rvert^2 - 2(\mathbf{u} \cdot \mathbf{v}) + \lvert\lvert\mathbf{v}\rvert\rvert^2$
$\lvert\lvert[-10, 5] - [3, 3]\rvert\rvert^2$	$\lvert\lvert[-10, 5]\rvert\rvert^2 - 2([-10, 5] \cdot [3, 3]) + \lvert\lvert[3, 3]\rvert\rvert^2$
$\lvert\lvert[-13, 2]\rvert\rvert^2$	$(100) + (25) - 2(-30 + 15) + (9) + (9)$
$169 + 4$	$125 + 30 + 18$
173	173

Norm Calculation Examples

- Suppose $\mathbf{u} = [-10, 5, 1]$ and $\mathbf{v} = [3, 3, 3]$. **Multiplication Norm Calculation**:

$\|\|\mathbf{u} \times \mathbf{v}\|\|^2$	$\|\|\mathbf{u}\|\|^2\|\|\mathbf{v}\|\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$
$\|\|[-10, 5, 1] \times [3, 3, 3]\|\|^2$	$\|\|[-10, 5, 1]\|\|^2\|\|[3, 3, 3]\|\|^2$ $-([-10, 5, 1] \cdot [3, 3, 3])^2$
$\|\|[(15) - (3), (3) - (-30), (-30) - (15)]\|\|^2$	$((100 + 25 + 1)(9 + 9 + 9)$ $-(-30 + 15 + 3)^2$
$(144) + (1089) + (2025)$	$(3402 - 144)$
3258	3258

Norming Vector Endpoints

- Norming can also be applied to find the n-dimensional distance between the endpoints of two vectors starting at the origin with a variant of the Pythagorean Theorem known as the *law of cosines*:

$$\|\mathbf{v}-\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\|\cos\theta,$$

  where θ is the angle from the $\mathbf{w}$ vector to the $\mathbf{v}$ vector measured in radians.

- This is also called the cosine rule and leads to the property that $\cos(\theta) = \frac{\mathbf{v}\mathbf{w}}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$.

Votes in the House of Commons

- Casstevens (1970) looked at legislative cohesion in the British House of Commons.
- Prime Minister David Lloyd George claimed on April 9, 1918 that the French Army was stronger on January 1, 1918 than on January 1, 1917 (a statement that generated considerable controversy).
- Subsequently the leader of the Liberal Party moved that a select committee be appointed to investigate claims by the military that George was incorrect.
- The resulting motion was defeated by the following vote: Liberal Party 98 yes, 71 no; Labour Party 9 yes, 15 no; Conservative Party 1 yes, 206 no; others 0 yes, 3 no.
- The difficult in analyzing this vote is the fact that 267 Members of Parliament (MPs) did not vote.
- So do we include them in the denominator when making claims about voting patterns?
- Casstevens says no because large numbers of abstentions mean that such indicators are misleading.

Votes in the House of Commons

- He alternatively looked at party cohesion for the two large parties as vector norms:

$$\|L\| = \|(98, 71)\| = 121.0165$$

$$\|C\| = \|(1, 206)\| = 206.0024.$$

- From this we get the obvious conclusion that the Conservatives are more cohesive because their vector has greater magnitude.
- More interestingly, we can contrast the two parties by calculating the angle between these two vectors (in radians) using the cosine rule:

$$\theta = \arccos \left[\frac{(98, 71) \cdot ((1, 206)}{121.070 \times 206.002} \right] = 0.9389,$$

  which is about 54 degrees.
- Recall that arccos is the inverse function to cos.

Vector p-Norm

- The norm described is the most commonly used form of a vector *p-norm*:

$$\|\mathbf{v}\|_p = (|v_1|^p + |v_2|^p + \cdots + |v_n|^p)^{\frac{1}{p}},\ p \geq 0,$$

  where $p = 2$ so far.

- Other important cases include $p = 1$ and $p = \infty$:

$$\|\mathbf{v}\|_\infty = \max_{1 \leq i \leq n} |x_i|,$$

  that is, just the maximum vector value.

- When a vector has a p-norm of 1, it is called a *unit vector*.

- If p is left off the norm, then one can safely assume that it is the $p = 2$ form.

Vector p-Norms Properties

Properties of Vector Norms, Length-n

$\rightarrowtail$ Triangle Inequality	$\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$
$\rightarrowtail$ Hölder's Inequality	for $\frac{1}{p}+\frac{1}{q}=1$, $\lvert\mathbf{v}\cdot\mathbf{w}\rvert \leq \|\mathbf{v}\|_p \|\mathbf{w}\|_q$
$\rightarrowtail$ Cauchy-Schwarz Ineq.	$\lvert\mathbf{v}\cdot\mathbf{w}\rvert \leq \|\mathbf{v}\|_2 \|\mathbf{w}\|_2$
$\rightarrowtail$ Cosine Rule	$\cos(\theta) = \frac{\mathbf{v}\mathbf{w}}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$
$\rightarrowtail$ Vector Distance	$d(\mathbf{v},\mathbf{w}) = \|\mathbf{v}-\mathbf{w}\|$
$\rightarrowtail$ Scalar Property	$\|s\mathbf{v}\| = \lvert s\rvert \|\mathbf{v}\|$

Hölder's Inequality Calculation

- Set $p = 3$ and $q = 3/2$ for the vectors $\mathbf{v} = [-1, 3]$ and $\mathbf{w} = [2, 2]$, respectively. Hölder's Inequality uses $|\mathbf{v} \cdot \mathbf{w}|$ to denote the absolute value of the dot product. Then:

$$\|\mathbf{v}\|_3 = (|-1|^3 + |3|^3)^{\frac{1}{3}} = 3.036589$$

$$\|\mathbf{w}\|_3 = (|2|^{\frac{3}{2}} + |2|^{\frac{3}{2}})^{\frac{2}{3}} = 3.174802$$

$$|\mathbf{v} \cdot \mathbf{w}| = |(-1)(2) + (3)(2)| = 4 < (3.036589)(3.174802) = 9.640569.$$

- So Hölder's Inequality barely holds here since the product of the first two quantities is 9.6406.

The Political Economy of Taxation

- While taxation is known to be an effective policy tool for democratic governments, it is also a very difficult political solution for many politicians because it can be unpopular and controversial.
- Swank and Steinmo (2002) looked at factors that lead to changes in tax policies in "advanced capitalist" democracies with the idea that factors like internationalization of economies, political pressure from budgets, and within-country economic factors are influential.
- They found that governments have a number of constraints on their ability to enact significant changes in tax rates, even when there is pressure to increase economic efficiency.
- As part of this study the authors provided a total taxation from labor and consumption as a percentage of GDP in the form of two vectors: one for 1981 and another for 1995. These are reproduced as

The Political Economy of Taxation

Nation	1981	1995
Australia	30	31
Austria	44	42
Belgium	45	46
Canada	35	37
Denmark	45	51
Finland	38	46
France	42	44
Germany	38	39
Ireland	33	34
Italy	31	41
Japan	26	29
Netherlands	44	44
New Zealand	34	38
Norway	49	46
Sweden	50	50
Switzerland	31	34
United Kingdom	36	36
United States	29	28

The Political Economy of Taxation

- A natural question to ask is, how much have taxation rates changed over the 14-year period for these countries collectively?
- The difference in mean averages, 38 versus 40, is not terribly revealing because it "washes out" important differences since some countries increased and other decreased.

The Political Economy of Taxation

- One way of assessing total country change is employing the difference norm, $||\mathbf{u}-\mathbf{v}||^2 = ||\mathbf{u}||^2 - 2(\mathbf{u}\cdot\mathbf{v}) + ||\mathbf{v}||^2$, to compare aggregate vector difference:

$$||t_{1995} - t_{1981}||^2 = t'_{1995}\cdot t_{1995} - 2(t'_{1995}\cdot t_{1981}) + t'_{1981}\cdot t_{1981}$$

$$= \begin{bmatrix} 30\\44\\45\\35\\45\\38\\42\\38\\33\\31\\26\\44\\34\\49\\50\\31\\36\\29 \end{bmatrix}' \cdot \begin{bmatrix} 30\\44\\45\\35\\45\\38\\42\\38\\33\\31\\26\\44\\34\\49\\50\\31\\36\\29 \end{bmatrix} - 2 \begin{bmatrix} 30\\44\\45\\35\\45\\38\\42\\38\\33\\31\\26\\44\\34\\49\\50\\31\\36\\29 \end{bmatrix}' \cdot \begin{bmatrix} 31\\42\\46\\37\\51\\46\\44\\39\\34\\41\\29\\44\\38\\46\\50\\34\\36\\28 \end{bmatrix} + \begin{bmatrix} 31\\42\\46\\37\\51\\46\\44\\39\\34\\41\\29\\44\\38\\46\\50\\34\\36\\28 \end{bmatrix}' \cdot \begin{bmatrix} 31\\42\\46\\37\\51\\46\\44\\39\\34\\41\\29\\44\\38\\46\\50\\34\\36\\28 \end{bmatrix}$$

$$= 260$$

The Political Economy of Taxation

- For comparison, we can calculate the same vector norm except that instead of using t_{1995}, we will substitute a vector that increases the 1981 uniformly levels by 10%
- A hypothetical increase of 10% for every country in the study:

$$\hat{t}_{1981} = 1.1t_{1981} = [33.0, 48.4, 49.5, 38.5, 49.5, 41.8, 46.2, 41.8, 36.3$$
$$34.1, 28.6, 48.4, 37.4, 53.9, 55.0, 34.1, 39.6, 31.9].$$

- This allows us to calculate the following benchmark difference:

$$||\hat{t}_{1981} - t_{1981}||^2 = 265.8.$$

- So now it is clear that the observed vector difference for total country change from 1981 to 1995 is actually similar to a 10% across-the-board change rather than a 5% change implied by the vector means. In this sense we get a true multidimensional sense of change.

Matrices

- A matrix is just a rectangular arrangement of numbers.
- It is a way to individually assign numbers, now called *matrix elements* or *entries*, to specified positions in a single structure, referred to with a single name.
- The order in which individual entries appear in the vector matters, the ordering of values within *both* rows and columns now matters.
- Matrices have two definable *dimensions*, the number of rows and the number of columns, denoted *row* × *column*.
- A matrix with i rows and j columns is said to be of dimension $i \times j$ (by convention rows comes before columns).
- A simple 2×2 matrix named $\mathbf{X}$ (like vectors, matrix names are bolded) is given by:

$$\underset{2\times 2}{\mathbf{X}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Matrices

- The numbers in a matrix are now treated together as a single unit.
- They are *grouped* together in the two-row by two-column matrix object (in the last example).
- The positioning of the numbers is specified, so that the previous matrix $\mathbf{X}$ is different than the following matrices:

$$\mathbf{W} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}, \qquad \mathbf{Z} = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix},$$

- The elements of a matrix can be integers, real numbers, or complex numbers.

Matrices

- We can refer directly to the specific elements of a matrix by using *subscripting* of addresses.
- The elements of the last $\mathbf{X}$ matrix are given by $x_{11} = 1$, $x_{12} = 2$, $x_{21} = 3$, and $x_{22} = 4$.
- The addresses of a $p \times n$ matrix can be described for large values of p and n by

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1(p-1)} & x_{1p} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2(p-1)} & x_{2p} \\ \vdots & \vdots & \ddots & & & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ x_{(n-1)1} & x_{(n-1)2} & \cdots & \cdots & x_{(n-1)(p-1)} & x_{(n-1)p} \\ x_{n1} & x_{n2} & \cdots & \cdots & x_{n(p-1)} & x_{np} \end{bmatrix}.$$

Matrix Equality

- Using subscripting notation we can now define *matrix equality*
- Matrix $\mathbf{A}$ is equal to matrix $\boldsymbol{B}$ if and only if every element of $\mathbf{A}$ is equal to the corresponding element of $\boldsymbol{B}$:
$$\boldsymbol{A} = \boldsymbol{B} \iff a_{ij} = b_{ij} \ \forall i, j.$$
- Note that subsumed in this definition is the requirement that the two matrices be of the same dimension (same number of rows, i, and columns, j).

Some Special Matrices

- A *square matrix* is a matrix with the same number of rows and columns.
- Because one number identifies the complete size of the square matrix, we can say that a $k \times k$ matrix is a matrix of **order-k**.
- Square matrices can contain any values and remain square: the square property is independent of the contents.
- A very general square matrix form is the *symmetric matrix*: symmetric across the diagonal from the upper left-hand corner to the lower right-hand corner.
- More formally, $\mathbf{X}$ is a symmetric matrix iff $a_{ij} = a_{ji}\ \forall i, j$.
- An example of a symmetric matrix:

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 8 & 5 & 6 \\ 3 & 5 & 1 & 7 \\ 4 & 6 & 7 & 8 \end{bmatrix}.$$

Some Special Matrices

- A matrix can also be *skew-symmetric* if it has the property that the rows and column switching operation would provide the same matrix except for the sign.
- For example,

$$\mathbf{X} = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix}.$$

- Note that the symmetric property does not hold for the *other* diagonal, the one from the upper right-hand corner to the lower left-hand corner.

Some Special Matrices

- Just as the symmetric matrix is a special case of the square matrix, the *diagonal matrix* is a special case of the symmetric matrix.
- A diagonal matrix is a symmetric matrix with all zeros on the off-diagonals (the values where $i \neq j$).
- If the (4×4) $\mathbf{X}$ matrix from before were a diagonal matrix, it would look like

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

Some Special Matrices

- We can also define the diagonal matrix more generally with just a vector.
- A diagonal matrix with elements $[d_1, d_2, \ldots, d_{n-1}, d_n]$ is the matrix

$$\mathbf{X} = \begin{bmatrix} d_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & d_{n-1} & 0 \\ 0 & 0 & 0 & 0 & d_n \end{bmatrix}.$$

- A diagonal matrix can have any values on the diagonal, but all of the other values must be zero.

The Identity Matrix

- A very important special case of the diagonal matrix is the *identity matrix*.
- This matrix has only the value 1 for each diagonal element: $d_i = 1, \ \forall i$.
- This *is* the matrix equivalent of the scalar number 1, pre- or post-multiplication of another (conformable) matrix gives the same matrix.
- A 4×4 version is

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

- This matrix form is always given the name $\mathbf{I}$, and it is sometimes denoted to give size: $I_{4\times4}$ or even just $\mathbf{I}(4)$.

The **J** Matrix

- A similar, but actually very different, matrix is the **J** matrix, which consists of all 1's.
- A 4×4 version is

$$\mathbf{J} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

- Note that this is not the matrix equivalent to the scalar 1.

The Zero Matrix

- The *zero matrix* is a matrix of all zeros.
- A 4×4 version is

$$\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

- Note that this *is* the matrix equivalent to the scalar 0: pre- or post-multiplication of another (conformable) matrix gives the zero matrix.

Triangular Matrices

- A special nonsymmetric square matrix is called the *triangular matrix*.
- This is a matrix with all zeros above the diagonal, *lower triangular*, or all zeros below the diagonal, *upper triangular*.
- Two versions based on the first square matrix given above are

$$\mathbf{X}_{LT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 8 & 0 & 0 \\ 3 & 5 & 1 & 0 \\ 4 & 6 & 7 & 8 \end{bmatrix}, \quad \mathbf{X}_{UT} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 8 & 5 & 6 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 8 \end{bmatrix},$$

Marriage Satisfaction

- Sociologists who study marriage often focus on indicators of self-expressed satisfaction.
- Unfortunately marital satisfaction is sufficiently complex and sufficiently multidimensional that single measurements are often insufficient to get a full picture of underlying attitudes.
- Norton (1983) asked multiple questions designed to elicit varied expressions of marital satisfaction and therefore care a lot about the correlation between these.
- A correlation shows how "tightly" two measures change with each other over a range from -1 to 1, with 0 being no evidence of moving together.
- His correlation matrix provides the correlational structure between answers to the following questions according to scales where higher numbers mean that the respondent agrees more (i.e., 1 is strong disagreement with the statement and 7 is strong agreement with the statement).

Marriage Satisfaction

- The questions are

Question	**Measurement Scale**	**Valid Cases**
We have a good marriage	7-point	428
My relationship with my partner is very stable	7-point	429
Our marriage is strong	7-point	429
My relationship with my partner makes me happy	7-point	429
I really feel like *part of a team* with my partner	7-point	426
The degree of happiness, everything considered	10-point	407

Marriage Satisfaction

- Since the correlation between two variables is symmetric, it does not make sense to give a correlation matrix between these variables across a full matrix because the lower triangle will simply mirror the upper triangle and make the display more congested.
- Norton only needs to show a triangular version of the matrix:

$$\begin{array}{c} \\ (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \end{array} \begin{array}{c} \begin{array}{cccccc} (1) & (2) & (3) & (4) & (5) & (6) \end{array} \\ \begin{pmatrix} 1.00 & 0.85 & 0.83 & 0.83 & 0.74 & 0.76 \\ & 1.00 & 0.82 & 0.86 & 0.72 & 0.77 \\ & & 1.00 & 0.78 & 0.68 & 0.70 \\ & & & 1.00 & 0.71 & 0.76 \\ & & & & 1.00 & 0.69 \\ & & & & & 1.00 \end{pmatrix} \end{array}.$$

Marriage Satisfaction

- These analyzed questions all correlate highly (a 1 means a perfectly positive relationship).
- The question that seems to covary greatly with the others is the first one.
- Notice that strength of marriage and part of a team covary less than any others (a suggestive finding).

Matrix Addition and Substraction

- **Matrix Addition**.

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} -2 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{X} + \mathbf{Y} = \begin{bmatrix} 1-2 & 2+2 \\ 3+0 & 4+1 \end{bmatrix} = \begin{bmatrix} -1 & 4 \\ 3 & 5 \end{bmatrix}.$$

- **Matrix Subtraction**.

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} -2 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{X} - \mathbf{Y} = \begin{bmatrix} 1-(-2) & 2-2 \\ 3-0 & 4-1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 3 & 3 \end{bmatrix}.$$

Matrix Scalar Multiplication and Division

- **Scalar Multiplication**.

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad s = 5$$

$$s \times \mathbf{X} = \begin{bmatrix} 5 \times 1 & 5 \times 2 \\ 5 \times 3 & 5 \times 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 15 & 20 \end{bmatrix}.$$

- **Scalar Division**.

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad s = 5$$

$$\mathbf{X} \div s = \begin{bmatrix} 1 \div 5 & 2 \div 5 \\ 3 \div 5 & 4 \div 5 \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix}.$$

Properties of Matrix Calculations

Properties of (Conformable) Matrix Manipulation

↣ Commutative Property	$\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$
↣ Additive Associative Property	$(\mathbf{X} + \mathbf{Y}) + \mathbf{Z} = \mathbf{X} + (\mathbf{Y} + \mathbf{Z})$
↣ Matrix Distributive Property	$s(\mathbf{X} + \mathbf{Y}) = s\mathbf{X} + s\mathbf{Y}$
↣ Scalar Distributive Property	$(s + t)\mathbf{X} = s\mathbf{X} + t\mathbf{X}$
↣ Zero Property	$\mathbf{X} + \mathbf{0} = \mathbf{X}$ and $\mathbf{X} - \mathbf{X} = \mathbf{0}$

Matrix Calculations

- This example illustrates several of the properties above where $s = 7$, $t = 2$, $\mathbf{X} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$, and $\mathbf{Y} = \begin{bmatrix} 3 & 4 \\ 0 & -1 \end{bmatrix}$. The left-hand side is

$$(s+t)(\mathbf{X}+\mathbf{Y}) = (7+2)\left(\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 4 \\ 0 & -1 \end{bmatrix}\right)$$
$$= 9\begin{bmatrix} 5 & 4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 45 & 36 \\ 9 & 0 \end{bmatrix},$$

and the right-hand side is

$$t\mathbf{Y} + s\mathbf{Y} + t\mathbf{X} + s\mathbf{X} = 2\begin{bmatrix} 3 & 4 \\ 0 & -1 \end{bmatrix} + 7\begin{bmatrix} 3 & 4 \\ 0 & -1 \end{bmatrix} + 2\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} + 7\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 6 & 8 \\ 0 & -2 \end{bmatrix} + \begin{bmatrix} 21 & 28 \\ 0 & -7 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 14 & 0 \\ 7 & 7 \end{bmatrix} = \begin{bmatrix} 45 & 36 \\ 9 & 0 \end{bmatrix}.$$

Matrix Multiplication

- Two matrices are *conformable* for multiplication if the number of columns in the first matrix match the number of rows in the second matrix.
- Note that this implies that *the order of multiplication matters with matrices.*
- This is the first algebraic principle that deviates from the simple scalar world.
- suppose that $\mathbf{X}$ is size $k \times n$ and $\mathbf{Y}$ is size $n \times p$. Then the multiplication operation given by

$$\underset{(k \times n)}{\mathbf{X}} \underset{(n \times p)}{\mathbf{Y}}$$

  is valid because the inner numbers match up.
- The multiplication operation given by

$$\underset{(n \times p)}{\mathbf{Y}} \underset{(k \times n)}{\mathbf{X}}$$

  is not unless $p = k$.

Matrix Multiplication

- The inner dimension numbers of the operation determine conformability and the outer dimension numbers determine the size of the resulting matrix.
- So in the example of $\mathbf{XY}$ above, the resulting matrix would be of size $k \times p$.
- To maintain awareness of this order of operation distinction, we say that $\mathbf{X}$ *pre-multiplies* $\mathbf{Y}$ or, equivalently, that $\mathbf{Y}$ *post-multiplies* $\mathbf{X}$.

Matrix Multiplication

- Consider matrix multiplication in *vector terms*.
- For $\mathbf{X}_{k\times n}$ and $\mathbf{Y}_{n\times p}$, we take each of the n row vectors in $\mathbf{X}$ and perform a vector inner product with the n column vectors in $\mathbf{Y}$.
- This operation starts with performing the inner product of the first row in $\mathbf{X}$ with the first column in $\mathbf{Y}$ and the result will be the first element of the product matrix:

$$\begin{aligned}\mathbf{XY} &= \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \\ &= \begin{bmatrix} (x_{11}\ x_{12})\cdot(y_{11}\ y_{21}) & (x_{11}\ x_{12})\cdot(y_{12}\ y_{22}) \\ (x_{21}\ x_{22})\cdot(y_{11}\ y_{21}) & (x_{21}\ x_{22})\cdot(y_{12}\ y_{22}) \end{bmatrix} \\ &= \begin{bmatrix} x_{11}y_{11}+x_{12}y_{21} & x_{11}y_{12}+x_{12}y_{22} \\ x_{21}y_{11}+x_{22}y_{21} & x_{21}y_{12}+x_{22}y_{22} \end{bmatrix}.\end{aligned}$$

Matrix Multiplication

- Suppose that we notate the four values of the final matrix as $\mathbf{XY}[1,1]$, $\mathbf{XY}[1,2]$, $\mathbf{XY}[2,1]$, $\mathbf{XY}[2,2]$ corresponding to their position in the 2×2 product.
- Then we can visualize how the rows of the first matrix operate against the columns of the second matrix to produce each value:

$$\begin{bmatrix} x_{11} & x_{12} \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{21} \end{bmatrix} = \mathbf{XY}[1,1], \quad \begin{bmatrix} x_{11} & x_{12} \end{bmatrix} \begin{bmatrix} y_{12} \\ y_{22} \end{bmatrix} = \mathbf{XY}[1,2],$$

$$\begin{bmatrix} x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{21} \end{bmatrix} = \mathbf{XY}[2,1], \quad \begin{bmatrix} x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} y_{12} \\ y_{22} \end{bmatrix} = \mathbf{XY}[2,2].$$

Matrix Multiplication in Scalar Notation

- scalar notation for an arbitrary-sized operation:

$$\underset{(k\times n)}{\mathbf{X}}\underset{(n\times p)}{\mathbf{Y}} = \begin{bmatrix} \sum_{i=1}^{n} x_{1i}y_{i1} & \sum_{i=1}^{n} x_{1i}y_{i2} & \cdots & \sum_{i=1}^{n} x_{1i}y_{ip} \\ \sum_{i=1}^{n} x_{2i}y_{i1} & \sum_{i=1}^{n} x_{2i}y_{i2} & \cdots & \sum_{i=1}^{n} x_{2i}y_{ip} \\ \vdots & & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ki}y_{i1} & \cdots & \cdots & \sum_{i=1}^{n} x_{ki}y_{ip} \end{bmatrix}.$$

- Now perform matrix multiplication with some actual values. Starting with the matrices

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} -2 & 2 \\ 0 & 1 \end{bmatrix},$$

calculate

$$\begin{aligned} \mathbf{XY} &= \begin{bmatrix} (1 \ \ 2)\cdot(-2 \ \ 0) & (1 \ \ 2)\cdot(2 \ \ 1) \\ (3 \ \ 4)\cdot(-2 \ \ 0) & (3 \ \ 4)\cdot(2 \ \ 1) \end{bmatrix} \\ &= \begin{bmatrix} (1)(-2)+(2)(0) & (1)(2)+(2)(1) \\ (3)(-2)+(4)(0) & (3)(2)+(4)(1) \end{bmatrix} \\ &= \begin{bmatrix} -2 & 4 \\ -6 & 10 \end{bmatrix}. \end{aligned}$$

Matrix Multiplication Properties

Properties of (Conformable) Matrix Multiplication

$\rightarrowtail$ Associative Property	$(\mathbf{XY})\mathbf{Z} = \mathbf{X}(\mathbf{YZ})$
$\rightarrowtail$ Additive Distributive Property	$(\mathbf{X} + \mathbf{Y})\mathbf{Z} = \mathbf{XZ} + \mathbf{YZ}$
$\rightarrowtail$ Scalar Distributive Property	$s\mathbf{XY} = (\mathbf{X}s)\mathbf{Y} = \mathbf{X}(s\mathbf{Y}) = \mathbf{XY}s$
$\rightarrowtail$ Zero Property	$\mathbf{X0} = \mathbf{0}$

LU Matrix Decomposition

- Many square matrices can be decomposed as the product of lower and upper triangular matrices.
- Given $\mathbf{A}$:
$$\underset{(p\times p)}{\mathbf{A}} = \underset{(p\times p)}{\mathbf{L}}\underset{(p\times p)}{\mathbf{U}},$$
where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{U}$ is an upper triangular matrix
- Consider the following example matrix decomposition according to this scheme:
$$\begin{bmatrix} 2 & 3 & 3 \\ 1 & 2 & 9 \\ 1 & 1 & 12 \end{bmatrix} = \begin{bmatrix} 1.0 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.5 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3.0 & 3.0 \\ 0 & 0.5 & 7.5 \\ 0 & 0.0 & 18.0 \end{bmatrix}.$$
- This decomposition is very useful for solving systems of equations because much of the mechanical work is already done by the triangularization.

Back To Square Matrices

- What about multiplying two square matrices?
- Both orders of multiplication are possible, but except for special cases the result will differ.
- For example using matrices $\mathbf{X}$ and $\mathbf{Y}$:

$$\mathbf{XY} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$$

$$\mathbf{YX} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}.$$

- Even in cases where pre-multiplication and post-multiplication are possible, these are different operations and matrix multiplication is not commutative.

Post-Multiplying by the Identity Matrix and the **J** Matrix

- For example, post-multiplying **X** with **I**:

$$\mathbf{XI} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} (1)(1)+(2)(0) & (1)(0)+(2)(1) \\ (3)(1)+(4)(0) & (3)(0)+(4)(1) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$

- And then with **J**:

$$\mathbf{XJ}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} (1)(1)+(2)(1) & (1)(1)+(2)(1) \\ (3)(1)+(4)(1) & (3)(1)+(4)(1) \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 7 & 7 \end{bmatrix}.$$

- Post-multiplying by **I** returns the **X** matrix to its original form, but post-multiplying by **J** produces a matrix where values are the sum by row.

Pre-Multiplying by the Identity Matrix and the J Matrix

- Pre-multiplying by **I** also returns the original matrix, but pre-multiplying by **J** gives

$$\mathbf{J}_2\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} (1)(1)+(1)(3) & (1)(2)+(1)(4) \\ (1)(1)+(1)(3) & (1)(2)+(1)(4) \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 4 & 6 \end{bmatrix},$$

- This is now the sum down columns assigned as row values.

- This means that the **J** matrix can be very useful in calculations (including linear regression methods).

- For the **J** matrix, particularly for nonsquare forms:

$$\underset{(p\times n)}{\mathbf{J}}\underset{(n\times k)}{\mathbf{J}} = n \underset{(p\times k)}{\mathbf{J}}.$$

Moving Rows and Columns of the Identity Matrix

- Suppose we wish to switch two rows of a specific matrix.
- To accomplish this we can multiply by an identity matrix where the placement of the 1 values is switched:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{31} & x_{32} & x_{33} \\ x_{21} & x_{22} & x_{23} \end{bmatrix}.$$

- This pre-multiplying matrix is called a *permutation matrix* because it permutes the matrix that it operates on.

Moving Rows and Columns of the Identity Matrix

- The effect of changing a single 1 value to some other scalar is fairly obvious:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ sx_{31} & sx_{32} & sx_{33} \end{bmatrix},$$

  but the effect of changing a single 0 value is not:

$$\begin{bmatrix} 1 & 0 & s \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

$$= \begin{bmatrix} x_{11} + sx_{31} & x_{12} + sx_{32} & x_{13} + sx_{33} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}.$$

Matrix Permutation Example

- Consider the following example of permutation with an off-diagonal nonzero value:

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 & 3 \\ 7 & 0 & 1 \\ 3 & 3 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} (1 \cdot 3 + 0 \cdot 7 + 3 \cdot 3) & (1 \cdot 2 + 0 \cdot 0 + 3 \cdot 3) & (1 \cdot 3 + 0 \cdot 1 + 3 \cdot 3) \\ (0 \cdot 3 + 0 \cdot 7 + 1 \cdot 3) & (0 \cdot 2 + 0 \cdot 0 + 1 \cdot 3) & (0 \cdot 3 + 0 \cdot 1 + 1 \cdot 3) \\ (0 \cdot 3 + 1 \cdot 7 + 0 \cdot 3) & (0 \cdot 2 + 1 \cdot 0 + 0 \cdot 3) & (0 \cdot 3 + 1 \cdot 1 + 0 \cdot 3) \end{bmatrix}$$

$$= \begin{bmatrix} 12 & 11 & 12 \\ 3 & 3 & 3 \\ 7 & 0 & 1 \end{bmatrix},$$

- This shows the switching of rows two and three as well as the confinement of multiplication by 3 to the first row.

Matrix Transposition

- The operation of *matrix transposition* is similar to what we did in the context of vectors: switching between column and row forms.
- For matrices, this is slightly more involved but straightforward to understand: simply switch rows and columns.
- The transpose of an $i \times j$ matrix $\mathbf{X}$ is the $j \times i$ matrix $\mathbf{X}'$, usually called "X prime" (sometimes denoted $\mathbf{X}^T$ though).
- For example,

$$\mathbf{X}' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}' = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

- In this way the inner structure of the matrix is preserved but the shape of the matrix is changed.
- Transposition allows us to calculate the "square" of some arbitrary-sized $i \times j$ matrix: $\mathbf{X}'\mathbf{X}$ is always conformable, as is $\mathbf{X}\mathbf{X}'$, even if $i \neq j$.

Properties of Matrix Transposition

Properties of Matrix Transposition

$\rightarrowtail$ Invertibility	$(\mathbf{X}')' = \mathbf{X}$
$\rightarrowtail$ Additive Property	$(\mathbf{X} + \mathbf{Y})' = \mathbf{X}' + \mathbf{Y}'$
$\rightarrowtail$ Multiplicative Property	$(\mathbf{X}\mathbf{Y})' = \mathbf{Y}'\mathbf{X}'$
$\rightarrowtail$ General Multiplicative Property	$(\mathbf{X}_1\mathbf{X}_2 \ldots \mathbf{X}_{n-1}\mathbf{X}_n)' = \mathbf{X}'_n\mathbf{X}'_{n-1} \ldots \mathbf{X}'_2\mathbf{X}'_1$
$\rightarrowtail$ Symmetric Matrix	$\mathbf{X}' = \mathbf{X}$
$\rightarrowtail$ Skew-Symmetric Matrix	$\mathbf{X} = -\mathbf{X}'$

Calculation With Matrix Transposition

- Suppose we have the three matrices:

$$\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 3 & 7 \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} -2 & -2 \\ 1 & 0 \end{bmatrix}.$$

- Then the following calculation of $(\mathbf{XY}'+\mathbf{Z})' = \mathbf{Z}'+\mathbf{YX}'$ illustrates the invertibility, additive, and multiplicative properties of transposition. The left-hand side is

$$(\mathbf{XY}'+\mathbf{Z})' = \left(\begin{bmatrix} 1 & 0 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix}' + \begin{bmatrix} -2 & -2 \\ 1 & 0 \end{bmatrix} \right)' = \left(\begin{bmatrix} 2 & 2 \\ 27 & 20 \end{bmatrix} + \begin{bmatrix} -2 & -2 \\ 1 & 0 \end{bmatrix} \right)' = \left(\begin{bmatrix} 0 & 0 \\ 28 & 20 \end{bmatrix} \right)',$$

and the right-hand side is

$$\mathbf{Z}' + \mathbf{YX}' = \begin{bmatrix} -2 & -2 \\ 1 & 0 \end{bmatrix}' + \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 7 \end{bmatrix}' = \begin{bmatrix} -2 & 1 \\ -2 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 27 \\ 2 & 20 \end{bmatrix} = \begin{bmatrix} 0 & 28 \\ 0 & 20 \end{bmatrix}.$$

Idemptotent Matrices

- An *idempotent matrix* is a matrix that has the multiplication property

$$\mathbf{XX} = \mathbf{X}^2 = \mathbf{X}$$

- And therefore the property

$$\mathbf{X}^n = \mathbf{XX} \cdots \mathbf{X} = \mathbf{X}, \quad n \in \mathcal{I}^+$$

  (i.e., n is some positive integer).

- The identity matrix and the zero matrix are idempotent, but there are others:

$$\begin{bmatrix} -1 & 1 & -1 \\ 2 & -2 & 2 \\ 4 & -4 & 4 \end{bmatrix}$$

Idemptotent Matrices

- If a matrix is idempotent, then the difference between this matrix and the identity matrix is also idempotent because

$$(\mathbf{I} - \mathbf{X})^2 = \mathbf{I}^2 - 2\mathbf{X} + \mathbf{X}^2 = \mathbf{I} - 2\mathbf{X} + \mathbf{X} = (\mathbf{I} - \mathbf{X}).$$

- Test this with the example matrix above:

$$(\mathbf{I} - \mathbf{X})^2 = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} -1 & 1 & -1 \\ 2 & -2 & 2 \\ 4 & -4 & 4 \end{bmatrix} \right)^2$$

$$= \begin{bmatrix} 2 & -1 & 1 \\ -2 & 3 & -2 \\ -4 & 4 & -3 \end{bmatrix}^2 = \begin{bmatrix} 2 & -1 & 1 \\ -2 & 3 & -2 \\ -4 & 4 & -3 \end{bmatrix}.$$

Nilpotent Matrix

- A square *nilpotent* matrix is one with the property that $\mathbf{X}^n = \mathbf{0}$, for a positive integer n.
- Clearly the zero matrix is nilpotent, but others exist as well.
- A basic 2×2 example is the nilpotent matrix

$$\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}.$$

Involutory Matrix

- A *involutory matrix*, has the property that when squared it produces an identity matrix.
- For example,

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^2 = \mathbf{I},$$

- Other more complicated forms obviously exist.

Vectorization of Matrices

- Occasionally it is convenient to rearrange a matrix into vector form.
- The most common way to do this is to "stack" vectors from the matrix on top of each other, beginning with the first column vector of the matrix, to form one long column vector.
- Specifically, to *vectorize* an $i \times j$ matrix $\mathbf{X}$, we consecutively stack the j-length column vectors to obtain a single vector of length ij.
- This is denoted $\text{vec}(\mathbf{X})$ and has some obvious properties, such as $s\text{vec}(\mathbf{X}) = \text{vec}(s\mathbf{X})$ for some vector s and $\text{vec}(\mathbf{X} + \mathbf{Y}) = \text{vec}(\mathbf{X}) + \text{vec}(\mathbf{Y})$ for matrices conformable by addition.
- Returning to the example matrix,

$$\text{vec}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}.$$

- It is not true that $\text{vec}(\mathbf{X}) = \text{vec}(\mathbf{X}')$ since the latter would stack rows instead of columns.

The Kronecker Product

- The *Kronecker product* (also called the *tensor product*), which comes up naturally in the statistical analyses of time series data (data recorded on the same measures of interest at different points in time).

- This is a slightly more abstract process but has the advantage that there is no conformability requirement.

- For the $i \times j$ matrix $\mathbf{X}$ and $k \times \ell$ matrix $\mathbf{Y}$, a Kronecker product is the $(ik) \times (j\ell)$ matrix

$$\mathbf{X} \otimes \mathbf{Y} = \begin{bmatrix} x_{11}\mathbf{Y} & x_{12}\mathbf{Y} & \cdots & \cdots & x_{1j}\mathbf{Y} \\ x_{21}\mathbf{Y} & x_{22}\mathbf{Y} & \cdots & \cdots & x_{2j}\mathbf{Y} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ x_{i1}\mathbf{Y} & x_{i2}\mathbf{Y} & \cdots & \cdots & x_{ij}\mathbf{Y} \end{bmatrix}.$$

- This is different than

$$\mathbf{Y} \otimes \mathbf{X} = \begin{bmatrix} y_{11}\mathbf{X} & y_{12}\mathbf{X} & \cdots & \cdots & y_{1j}\mathbf{X} \\ y_{21}\mathbf{X} & y_{22}\mathbf{X} & \cdots & \cdots & y_{2j}\mathbf{X} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ y_{i1}\mathbf{X} & y_{i2}\mathbf{X} & \cdots & \cdots & y_{1j}\mathbf{X} \end{bmatrix}.$$

The Kronecker Product

- A numerical example of a Kronecker product follows for a (2×2) by (2×3) case:

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} -2 & 2 & 3 \\ 0 & 1 & 3 \end{bmatrix}$$

$$\mathbf{X} \otimes \mathbf{Y} = \begin{bmatrix} 1 \begin{bmatrix} -2 & 2 & 3 \\ 0 & 1 & 3 \end{bmatrix} & 2 \begin{bmatrix} -2 & 2 & 3 \\ 0 & 1 & 3 \end{bmatrix} \\ 3 \begin{bmatrix} -2 & 2 & 3 \\ 0 & 1 & 3 \end{bmatrix} & 4 \begin{bmatrix} -2 & 2 & 3 \\ 0 & 1 & 3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -2 & 2 & 3 & -4 & 4 & 6 \\ 0 & 1 & 3 & 0 & 2 & 6 \\ -6 & 6 & 9 & -8 & 8 & 12 \\ 0 & 3 & 9 & 0 & 4 & 12 \end{bmatrix},$$

The Kronecker Product

- which is different from the operation performed in reverse order:

$$\mathbf{Y} \otimes \mathbf{X} = \begin{bmatrix} -2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & 3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ 0\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & 1\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & 3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} -2 & -4 & 2 & 4 & 3 & 6 \\ -6 & -8 & 6 & 8 & 9 & 12 \\ 0 & 0 & 1 & 2 & 3 & 6 \\ 0 & 0 & 3 & 4 & 9 & 12 \end{bmatrix},$$

even though the resulting matrices are of the same dimension.

Kronecker Product Properties

Properties of Kronecker Products

$\rightarrowtail$ Trace $\quad \mathrm{tr}(\mathbf{X} \otimes \mathbf{Y}) = tr\mathbf{X} \otimes tr\mathbf{Y}$

$\rightarrowtail$ Transpose $\quad (\mathbf{X} \otimes \mathbf{Y})' = \mathbf{X}' \otimes \mathbf{Y}'$

$\rightarrowtail$ Inversion $\quad (\mathbf{X} \otimes \mathbf{Y})^{-1} = \mathbf{X}^{-1} \otimes \mathbf{Y}^{-1}$

$\rightarrowtail$ Products $\quad (\mathbf{X} \otimes \mathbf{Y})(\mathbf{W} \otimes \mathbf{Z}) = \mathbf{XW} \otimes \mathbf{YZ}$

$\rightarrowtail$ Associative $\quad (\mathbf{X} \otimes \mathbf{Y}) \otimes \mathbf{W} = \mathbf{X} \otimes (\mathbf{Y} \otimes \mathbf{W})$

$\rightarrowtail$ Distributive $\quad (\mathbf{X} + \mathbf{Y}) \otimes \mathbf{W} = (\mathbf{X} \otimes \mathbf{W}) + (\mathbf{Y} \otimes \mathbf{W})$

Example Calculations of Kronecker Product Properties

- Given the following matrices:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 2 & 5 \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} -1 & -3 \\ 1 & 1 \end{bmatrix} \quad \mathbf{W} = \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix},$$

we demonstrate that $(\mathbf{X} + \mathbf{Y}) \otimes \mathbf{W} = (\mathbf{X} \otimes \mathbf{W}) + (\mathbf{X} \otimes \mathbf{W})$.

- The left-hand side is

$$(\mathbf{X} + \mathbf{Y}) \otimes \mathbf{W} = \left(\begin{bmatrix} 1 & 1 \\ 2 & 5 \end{bmatrix} + \begin{bmatrix} -1 & -3 \\ 1 & 1 \end{bmatrix} \right) \otimes \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 3 & 6 \end{bmatrix} \otimes \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} & -2 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \\ 3 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} & 6 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 0 & -4 & 4 \\ 0 & 0 & -6 & 0 \\ 6 & -6 & 12 & -12 \\ 9 & 0 & 18 & 0 \end{bmatrix},$$

Kronecker Product Properties

- The right-hand side, $(\mathbf{X} \otimes \mathbf{W}) + (\mathbf{X} \otimes \mathbf{W})$, is

$$= \left(\begin{bmatrix} 1 & 1 \\ 2 & 5 \end{bmatrix} \otimes \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \right) + \left(\begin{bmatrix} -1 & -3 \\ 1 & 1 \end{bmatrix} \otimes \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 1 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} & 1 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \\ 2 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} & 5 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} -1 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} & -3 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \\ 1 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} & 1 \begin{bmatrix} 2 & -2 \\ 3 & 0 \end{bmatrix} \end{bmatrix},$$

- This simplifies down to

$$= \begin{bmatrix} 2 & -2 & 2 & -2 \\ 3 & 0 & 3 & 0 \\ 4 & -4 & 10 & -10 \\ 6 & 0 & 15 & 0 \end{bmatrix} + \begin{bmatrix} -2 & 2 & -6 & 6 \\ -3 & 0 & -9 & 0 \\ 2 & -2 & 2 & -2 \\ 3 & 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -4 & 4 \\ 0 & 0 & -6 & 0 \\ 6 & -6 & 12 & -12 \\ 9 & 0 & 18 & 0 \end{bmatrix}.$$