Linear Algebra Primer

Agenda Matrix-vector notation, solving matrix-vector equations, matrix multiplication, matrix inverse, transpose

HW
- practice matrix multiplication
- learn how to invert a matrix online

Matrix-Vector Notation

In 8^{th} grade (?) you studied <u>algebra</u> which is the manipulation of equations with symbols

$$4+3x=2$$

$$\Rightarrow x=-\frac{3}{2}$$

In statistics we often use <u>linear algebra</u> which is the manipulation of systems of equations with linear terms (e.g. x but no $x^2, x^3, x^4, \ldots$ terms),

$$\begin{cases} x_1+x_2=0 \\ x_1-x_2=0 \end{cases}$$

$$\Rightarrow x_1=0,\ x_2=0$$

But we usually use β (beta) as our symbols in statistics,

$$\begin{cases} \beta_1 + \beta_2 = 0 \\ \beta_1 - \beta_2 = 0 \end{cases}$$

$$\Longrightarrow \beta_1 = 0, \ \beta_2 = 0$$

Another way to write a system of linear equations is in <u>matrix-vector form</u>,

so
$$\begin{cases} \beta_1 + \beta_2 = 0 \\ \beta_1 - \beta_2 = 0 \end{cases}$$

is equivalent to
$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{2\times 2 \text{ matrix}} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

2x1 vectors

A <u>matrix</u> is a $n \times p$ grid of numbers,
n = num. rows, p = num columns,
a <u>column vector</u> is a $n \times 1$ matrix (column).

Often times in statistics, we deal with huge $n \times p$ matrices with large n and p, so we simply call the matrix X,

where $X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & \cdots & & X_{2p} \\ \vdots & & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$

The vector of variables is $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$

and the Right-hand-side vector, called the Response column/outcome vector, is $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

So, a system of linear equations like

$$\begin{cases} X_{11} \cdot \beta_1 + X_{12} \cdot \beta_2 + \dots + X_{1p} \cdot \beta_p = y_1 \\ X_{21} \cdot \beta_1 + X_{22} \cdot \beta_2 + \dots + X_{2p} \cdot \beta_p = y_2 \\ \vdots \\ X_{n1} \cdot \beta_1 + X_{n2} \cdot \beta_2 + \dots + X_{np} \cdot \beta_p = y_p \end{cases}$$

can be written in matrix-vector form as

$$\underbrace{\underbrace{\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & & & \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}}_{\text{matrix}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}}_{\text{vector}}}_{\text{matrix-vector multiplication}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

and can be written more concisely as

$$X\beta = y.$$

X is a matrix, X_{ij} is the number in X at row i and column j

β is a vector, β_i is the number in β at row i

y is a vector too.

Often times in statistics, X is a matrix of Known numbers, y is a vector of Known numbers, and β is a vector of unknown numbers.

Often we want to find β which satisfies the equation $X\beta = y$, which as you may recall is a linear system of equations.

Solving A System of Linear Equations, i.e. a Matrix-Vector Equation

- When is a matrix-vector equation solvable?
- If so, how to solve it?

Recall a simple 8^{th} grade algebra equation

$$3x = 2.$$

To solve this, you multiplied both sides by $\frac{1}{3}$, or the multiplicative inverse of 3,

$$3^{-1} \cdot 3x = 3^{-1} \cdot 2$$

$$\frac{1}{3} \cdot 3x = \frac{1}{3} \cdot 2$$

$$x = 2/3.$$

Now if we have matrix-vector equation

$$X \cdot \beta = y$$

where X, y is known, and we'd like to solve for the unknown β vector, it would be great if we could simply invert X,

$$X^{-1} \cdot X \beta = X^{-1} \cdot y$$

$$I \cdot \beta = X^{-1} y$$

$$\beta = X^{-1} y$$

I is the multiplicative identity matrix, so $I \cdot \beta = \beta$ and $\beta \cdot I = \beta$ for all β, similar to how in regular algebra $1 \cdot x = x \cdot 1 = x$.

$$\underset{p\times p}{I} = \begin{bmatrix} 1 & 0 & 0 & & \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & \\ & & & \ddots & \\ & 0 & & & 1 \end{bmatrix} = \begin{array}{l} \text{all 1's on diagonal} \\ \text{all 0's everywhere else} \end{array}$$

because

$$I \cdot \beta = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} 1\cdot\beta_1 + 0\cdot\beta_2 + \dots + 0\cdot\beta_p \\ 0\cdot\beta_1 + 0\cdot\beta_2 + \dots + 0\cdot\beta_p \\ \vdots \\ 0\cdot\beta_1 + 0\cdot\beta_2 + \dots + 0\cdot\beta_p \end{bmatrix}$$

$$= \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$$

and same for $\beta \cdot I = \beta$.

- Can we do this $X^{-1} \cdot X = I$??
- What does it even mean to multiply matrices like $X^{-1} X$??

Matrix-Vector Multiplication

Recall

$$\begin{cases} X_{11}\cdot\beta_1 + X_{12}\cdot\beta_2 + \cdots + X_{1p}\cdot\beta_p - \\ X_{21}\cdot\beta_1 + X_{22}\cdot\beta_2 + \cdots + X_{2p}\cdot\beta_p = \\ \vdots \\ X_{n1}\cdot\beta_1 + X_{n2}\cdot\beta_2 + \cdots + X_{np}\cdot\beta_p \end{cases}$$

is equivalent to

$$\underbrace{\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & & & \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}}_{\text{matrix-vector multiplication}}$$

For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 9 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1\cdot1 + 2\cdot2 + 3\cdot3 \\ 9\cdot1 + 1\cdot2 + 1\cdot3 \\ 1\cdot1 + 1\cdot2 + 1\cdot3 \end{bmatrix} = \begin{bmatrix} 14 \\ 14 \\ 6 \end{bmatrix}$$

Matrix multiplication is like multiple matrix-vector multiplications, for example

$$\begin{bmatrix} 1 & 2 & 3 \\ 9 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 14 & 17 \\ 14 & 15 \\ 6 & 7 \end{bmatrix}$$

So, for 2 matrices A and B,

$$\underset{n_A \times p_A}{A} \cdot \underset{n_B \times p_B}{B}$$

also, $\underbrace{A \cdot B}_{n_A \times p_B}$

only makes sense if $p_A = n_B$.

So, with $X \cdot \beta = y$

we wanted matrix inverse X^{-1} so that

$$\underset{a \times b}{X^{-1}} \underset{n \times p}{X} \underset{p \times 1}{\beta} = \underset{a \times b}{X^{-1}} \underset{n \times 1}{y}$$

So we need $b = n$.

much like $3x = 2$, we want 3^{-1} so that

$3^{-1} \cdot 3x = 3^{-1} \cdot 2 \Rightarrow \frac{1}{3} \cdot 3x = \frac{1}{3} \cdot 2 \Rightarrow x = \frac{2}{3}$

But $3 \cdot 3^{-1} = 3 \cdot \frac{1}{3} = 1$ also,

So, $3^{-1} = \frac{1}{3}$ is a <u>left inverse</u> and <u>Right inverse</u>,

So, we also want $\underset{n \times p}{X} \underset{a \times b}{X^{-1}}$, so we need $a = p$.

But for $\underbrace{\underset{n\times p}{X}\underset{p\times n}{X^{-1}}}_{n\times n} = \underbrace{\underset{p\times n}{X^{-1}}\underset{n\times p}{X}}_{p\times p}$, we need $n=p$.

So, $\boxed{\text{invertible matrices are square}}$ $p\times p$.

For example, if $X=\begin{bmatrix}1&0\\0&2\end{bmatrix}$, $X^{-1}=\begin{bmatrix}1&0\\0&1/2\end{bmatrix}$

because $XX^{-1}=\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&0\\0&1/2\end{bmatrix} = \begin{pmatrix}1\cdot1+0\cdot0 & 1\cdot0+0\cdot\frac{1}{2}\\ 0\cdot1+2\cdot0 & 0\cdot0+2\cdot\frac{1}{2}\end{pmatrix}$

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

and $X^{-1}X = \begin{bmatrix}1&0\\0&1/2\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix} = \begin{bmatrix}1&0\\0&1\end{bmatrix} = I$.

There are good algorithms to invert a general square matrix, if it is possible to do so.

I refer you to google, or Math 2400, to learn more.

For now, assume you can invert (many) square matrices with ease.

$\boxed{\text{\underline{HW} learn how to invert a matrix, online}}$

Often times in statistics, we have matrix-vector equations

$$\underset{n\times p}{X} \cdot \underset{p\times 1}{\beta} = \underset{n\times 1}{y}$$

where $n \neq p$. Usually, $n >> p$ in sports.

Sometimes $p > n$.

Usually, n = number of data examples,
p = number of "features".

For example, n = # games played
p = # teams
to make regression based power ratings.

So, we can't invert X since X is not square.

Fortunately, we can use a trick to make a square matrix appear, and then we can invert it!

The <u>transpose</u> of a matrix X, written X^T, simply involves switching the rows and columns, for example,

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \Longrightarrow X^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

So the transpose of X is defined by $\boxed{X^T_{ij} = X_{ji}}$

and

$$X^TX = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 1\cdot1+3\cdot3+5\cdot5 & 1\cdot2+3\cdot4\cdot5\cdot6 \\ 2\cdot1+4\cdot3+6\cdot5 & 2\cdot2+4\cdot4+6\cdot6 \end{bmatrix}$$

$$= \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix}.$$

So, $\boxed{\underset{p\times p}{X^TX} \text{ is square and symmetric}}$

A <u>symmetric</u> matrix A has $A_{ij} = A_{ji}$

$$A = \begin{pmatrix} 35 & 44 \\ 44 & 56 \end{pmatrix} \qquad A_{12} = 44 = A_{21}$$

$X^T X$ is symmetric because

$$(X^T X)_{ij} = \begin{pmatrix} i^{th} \text{ Row} \\ \text{of } X^T \end{pmatrix} \bullet \begin{pmatrix} j^{th} \text{ column} \\ \text{of } X \end{pmatrix}$$

$$= \begin{pmatrix} i^{th} \text{ col} \\ \text{of } X \end{pmatrix} \bullet \begin{pmatrix} j^{th} \text{ Row} \\ \text{of } X^T \end{pmatrix}$$

$$= \begin{pmatrix} j^{th} \text{ Row} \\ \text{of } X^T \end{pmatrix} \bullet \begin{pmatrix} i^{th} \text{ col} \\ \text{of } X \end{pmatrix}$$

$$= (X^T X)_{ji}$$

$\square$

Because $\underset{p \times p}{X^T X}$ is square and symmetric, it is (usually) easy to invert it, and so a common statistical equation solving process looks like

$$X\beta = y$$

$$X^T X \beta = X^T y$$

$$(X^T X)^{-1}(X^T X)\beta = (X^T X)^{-1} X^T y$$

$$\beta = (X^T X)^{-1} X^T y.$$

- But still, how to invert a square (and, symmetric) matrix ??

To be continued...

What it means to do Research

Research is a series of incremental improvements in answering a question.

1. Ask a question
   - What's something I want to know?
   - Read (e.g. Neil Paine, blogs, Twitter, JQAS)
2. Is it a good research question for me today?
   - right level of difficulty?
   - would I potentially be able to answer it?
   - has it been done before?
   - is it quantitative?
   - will anyone care?
   - will I care?
   - do I have the passion to put in the time commitment to try to answer it?
3. Literature Review
   - What's been done?
   - Related work
   - Brush up on prerequisites
   - Find your starting point

4. From English to Math/Code/Modeling

e.g. Idea in English: game-by-game wins above replacement for starting pitchers

Math: come up with the formulas

Code: implement formular with data