CSSS Math Camp Lecture 2

Matrix Algebra

Authored by: Laina Mercer, PhD

Erin Lipman and Jess Kunke

Department of Statistics

September 12, 2023

Lecture 2: Matrix Algebra

- Matrix Algebra
  - Definitions, notation
  - Matrix Arithmetic
  - Determinants - existence of an inverse
  - Linear equations
  - Least Squares and Regression with matrices

Motivation

Matrix algebra provides concise notation and rules for manipulating matrices (arrays of numbers).

Matrix algebra will be important for computing linear regression estimates.

Motivation

Example dataframe:

```
  region years       u5m     lower     upper
1    All 80-84 0.1691030 0.1573394 0.1815566
2    All 85-89 0.1603335 0.1490694 0.1722763
3    All 90-94 0.1208087 0.1079371 0.1349829
4  tanga 80-84 0.1810487 0.1369700 0.2354425
5  tanga 85-89 0.2230574 0.1677716 0.2902086
```

Definitions & Notation

What is a matrix?

A *matrix* is an array of number is a rectangular form.
Examples:

$$A = \begin{bmatrix} 1 & 2 & 6 & 4 \\ 5 & 8 & 12 & 8 \\ 4 & 3 & 2 & 1 \end{bmatrix} B = \begin{bmatrix} 4 & 3 & 2 \\ 1 & 2 & 4 \end{bmatrix}$$

where A is a 3×4 matrix and B is a 2×3 matrix. Note: matrix dimensions are always listed as rows $\times$ columns.

Definitions & Notation

What is a matrix?

In mathematical notation, a matrix is written

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

Where x_{ij} is the value in the ith row and the jth column of matrix X.

A *vector* is a matrix that has n rows and 1 column (or 1 row and n columns).

Examples:

$$\begin{bmatrix} 1 & 2 & 6 & 4 \end{bmatrix} \text{ or } \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix}$$

Definitions & Notation

Special Matrices

A *vector* is a matrix that has n rows and 1 column (or 1 row and n columns).
Examples:

$$\begin{bmatrix} 1 & 2 & 6 & 4 \end{bmatrix} \text{ or } \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix}$$

A *square* matrix has the same number of rows and columns.
Example:

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

A *symmetric* matrix has elements such that $x_{ij} = x_{ji}$.
Example:

$$\begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 7 \end{bmatrix}$$

A *symmetric* matrix has elements such that $x_{ij} = x_{ji}$.
Example:

$$\begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 7 \end{bmatrix}$$

A symmetric matrix must also be a square matrix.

Definitions & Notation

Special Matrices

A *diagonal* matrix is a matrix that is zero everywhere except on the diagonal. Where the diagonal is defined as all elements for which the row number is equal to the column number $\{(1,1),(2,2),(3,3),...\}$.

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

Definitions & Notation

Special Matrices

A *diagonal* matrix is a matrix that is zero everywhere except on the diagonal. Where the diagonal is defined as all elements for which the row number is equal to the column number $\{(1,1),(2,2),(3,3),...\}$.

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

A special case of a diagonal matrix is the *identity* matrix. Its diagonal elements are all ones.

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

Clearly, the identity matrix (or any other diagonal matrix) is also symmetric.

Matrix Equality: Two matrices A, B are equal if and only if, for all elements, each $a_{ij} = b_{ij}$. (Note: this means they must have the same dimensions.)

Matrix Transpose: The *transpose* of a matrix is found by interchanging the corresponding rows and columns of a matrix. The first row becomes the first column, the second row becomes the second clump, etc. The dimensions are then switched and the element a_{ij} becomes the element a_{ji}. The transposed matrix is often denoted A^t (or A').

Matrix Equality: Two matrices A, B are equal if and only if, for all elements, each $a_{ij} = b_{ij}$. (Note: this means they must have the same dimensions.)

Matrix Transpose: The *transpose* of a matrix is found by interchanging the corresponding rows and columns of a matrix. The first row becomes the first column, the second row becomes the second clump, etc. The dimensions are then switched and the element a_{ij} becomes the element a_{ji}. The transposed matrix is often denoted A^t (or A').

$$A = \begin{bmatrix} 1 & 2 & 6 \\ 3 & 5 & 9 \end{bmatrix} \quad A^t = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 6 & 9 \end{bmatrix}$$

Matrix Arithmetic

Addition & Subtraction

Two matrices can be added or subtracted only if their dimensions are the same (both rows and columns). The corresponding elements are then added or subtracted.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

Two matrices can be added or subtracted only if their dimensions are the same (both rows and columns). The corresponding elements are then added or subtracted.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}$$

Example:

$$\begin{bmatrix} 1 & 2 & 6 \\ 3 & 5 & 9 \end{bmatrix} - \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -2 \\ -3 & -4 & 3 \end{bmatrix}$$

Matrix Arithmetic

Scaler Multiplication

To multiply a matrix by a *scalar* (a constant value), multiply each element by that number.
Example:

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \end{bmatrix} \quad 3A = \begin{bmatrix} 3 & 9 & 24 \\ 18 & 27 & 18 \end{bmatrix}$$

Matrix Multiplication

Two matrices can be multiplied only if the number of columns of the first matrix equals the number of rows of the second matrix (the inside numbers). The resulting matrix has dimensions: number of rows of first matrix by number of columns of second matrix (the outside numbers).

Matrix Arithmetic

Matrix Multiplication

Two matrices can be multiplied only if the number of columns of the first matrix equals the number of rows of the second matrix (the inside numbers). The resulting matrix has dimensions: number of rows of first matrix by number of columns of second matrix (the outside numbers).

Multipling a row vector by a column vector:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$$

A is (1×3); B is (3×1), AB is (1×1), i.e. a scalar.

Matrix Arithmetic

Matrix Multiplication

Two matrices can be multiplied only if the number of columns of the first matrix equals the number of rows of the second matrix (the inside numbers). The resulting matrix has dimensions: number of rows of first matrix by number of columns of second matrix (the outside numbers).

Multipling a row vector by a column vector:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$$

A is (1×3); B is (3×1), AB is (1×1), i.e. a scalar.

We multiply a row and column vector by taking the cross product of the numbers in A and B.

$$A \cdot B = a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31}$$

Matrix Arithmetic

Matrix Multiplication

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

A is (2×3); B is $(3\times\ 2)$.
$A \cdot B$ is $(2 \times 3 \cdot 3 \times 2)$ results in $2{\times}2$.
$B \cdot A$ is $(3 \times 2 \cdot 2 \times 3)$ results in $3{\times}3$.

Matrix Arithmetic

Matrix Multiplication

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

A is (2×3); B is $(3\times\ 2)$.
$A \cdot B$ is $(2 \times 3 \cdot 3 \times 2)$ results in $2{\times}2$.
$B \cdot A$ is $(3 \times 2 \cdot 2 \times 3)$ results in $3{\times}3$.

We find each element $(ab)_{ij}$ by summing the crossproducts of the ith row of A and the jth column of B.

$$A{\cdot}B = \begin{bmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} & a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + a_{13} \cdot b_{32} \\ a_{21} \cdot b_{21} + a_{12} \cdot b_{21} + a_{23} \cdot b_{31} & a_{21} \cdot b_{12} + a_{22} \cdot b_{22} + a_{23} \cdot b_{32} \end{bmatrix}$$

Matrix Arithmetic

Multiplication Examples

Examples:

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \\ 2 & 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 9 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} 1\cdot 3 + 3\cdot 2 + 8\cdot 3 & 1\cdot 9 + 3\cdot 1 + 8\cdot 2 \\ 6\cdot 3 + 9\cdot 2 + 6\cdot 3 & 6\cdot 9 + 9\cdot 1 + 6\cdot 2 \\ 2\cdot 3 + 1\cdot 2 + 3\cdot 3 & 2\cdot 9 + 1\cdot 1 + 3\cdot 2 \end{bmatrix} = \begin{bmatrix} 33 & 28 \\ 54 & 75 \\ 17 & 25 \end{bmatrix}$$

Examples:

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \\ 2 & 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 9 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} 1\cdot 3 + 3\cdot 2 + 8\cdot 3 & 1\cdot 9 + 3\cdot 1 + 8\cdot 2 \\ 6\cdot 3 + 9\cdot 2 + 6\cdot 3 & 6\cdot 9 + 9\cdot 1 + 6\cdot 2 \\ 2\cdot 3 + 1\cdot 2 + 3\cdot 3 & 2\cdot 9 + 1\cdot 1 + 3\cdot 2 \end{bmatrix} = \begin{bmatrix} 33 & 28 \\ 54 & 75 \\ 17 & 25 \end{bmatrix}$$

Note: $A \cdot B$ is not necessarily equal to $B \cdot A$. For matrix multiplication, order matters. In this case $B \cdot A$ cannot be computed as the dimensions are not compatible ($3 \times 2 \cdot 3 \times 3$).

Matrix Arithmetic

Practice

What is the dimension of AB? Is it possible to compute BA? Find AB.

$$A = \begin{bmatrix} 3 & 1 & 4 & 1 \\ 5 & 9 & 2 & 6 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Matrix Inverse:

The inverse of a number is its reciprocal; a number multiplied by its inverse equals 1. $(4 \cdot 1/4 = 1)$

Matrix Inverse:

The inverse of a number is its reciprocal; a number multiplied by its inverse equals 1. $(4 \cdot 1/4 = 1)$

The inverse of a matrix A is the matrix A^{-1} that satisfies $A \cdot A^{-1} = I$. Where I is the identity matrix (ones along the diagonal and the rest are zeros). **Only square matrices can have inverses.**

Matrix Inverse:

The inverse of a number is its reciprocal; a number multiplied by its inverse equals 1. $(4 \cdot 1/4 = 1)$

The inverse of a matrix A is the matrix A^{-1} that satisfies $A \cdot A^{-1} = I$. Where I is the identity matrix (ones along the diagonal and the rest are zeros). **Only square matrices can have inverses.**

Remember that matrix multiplication is not just multiplying pairs of elements, so we can't just find the reciprocal of each element. So how do we find the inverse? How do we know if the inverse exists?

The *determinant* is a number that can be computed for any square matrix.

For a 2×2 matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the determinant is $D(A) = a \cdot d - b \cdot c$.

The *determinant* is a number that can be computed for any square matrix.

For a 2×2 matrix,

$$A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$$

the determinant is $D(A) = a \cdot d - b \cdot c$.

If the determinant is zero, no inverse exists. If the determinant is nonzero then the inverse exists.

If no nverse exists the matrix is called *singular*.

Matrix Arithmetic

Determinant Example

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}$$

$D = 4 \cdot 6 - 12 \cdot 3 = -12$. Inverse exists.

Matrix Arithmetic

Determinant Example

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}$$

$D = 4 \cdot 6 - 12 \cdot 3 = -12$. Inverse exists.

$$A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

$D = 2 \cdot 2 - 4 \cdot 1 = 0$. Inverse does not exist,

Once we know the inverse exists, we can find it.
For a 2×2 matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{D(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

where again $D(A) = a \cdot d - b \cdot c$. Example:

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}, \quad A^{-1} = \frac{1}{-12} \begin{bmatrix} 6 & -12 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -1/2 & 1 \\ 1/4 & -1/3 \end{bmatrix}$$

Note: Finding the inverse for higher dimensions involves more comp;located formulas and is usually solved by a math software.

Linear Equations

Let's go back to thinking about systems of two equations:

$$ax + by = g$$

$$cx + dy = f$$

Previously we solved this system by eliminating the y variable, solving for x, and then substituting back in for y.

No we can write this system in matrix notation:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} g \\ f \end{bmatrix}$$

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

So how do we solve for z?

Linear Equations

Examples

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

So how do we solve for z? First, left-multiply the equation by A^{-1}:

$$A^{-1} \cdot A \cdot z = A^{-1} \cdot w$$

By definition $A^{-1} \cdot A = I$. Thus,

$$I \cdot z = A^{-1} \cdot w \text{ or } z = A^{-1} \cdot w.$$

So we can find $z = (x, y)$, the solution to our system, by finding $z = A^{-1} \cdot w$.

$$2x + y = 1$$
$$4x + 3y = 8$$

$$2x + y = 1$$

$$4x + 3y = 8$$

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

Linear Equations

Examples

$$2x + y = 1$$

$$4x + 3y = 8$$

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

$$A^{-1} = \frac{1}{2 \cdot 3 - 4 \cdot 1} \begin{bmatrix} 3 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix}$$

Linear Equations

Examples

$$2x + y = 1$$

$$4x + 3y = 8$$

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

$$A^{-1} = \frac{1}{2 \cdot 3 - 4 \cdot 1} \begin{bmatrix} 3 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix}$$

$$z = A^{-1} \cdot w = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 8 \end{bmatrix} = \begin{bmatrix} 3/2 \cdot 1 + -1/2 \cdot 8 \\ -2 \cdot 1 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 6 \end{bmatrix}$$

Least Squares

Often we have much more data that pairs of points. We may have a survey where we asked n people the same p questions. We can put that data in a matrix of dimensions $n \times p$, where each row is a person and each column is one of the asked questions.

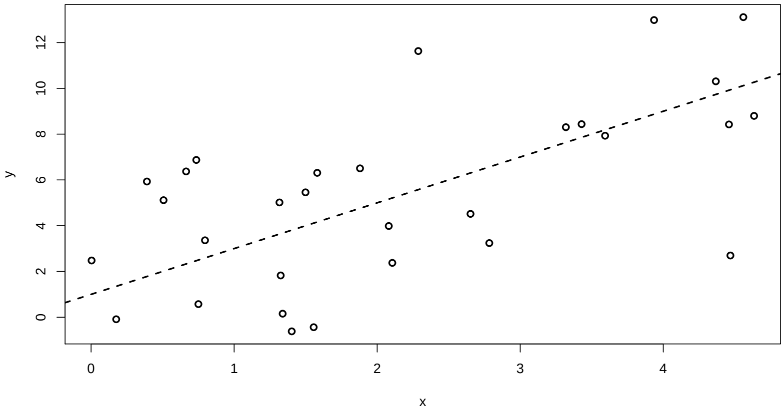
Least Squares

Often we have much more data that pairs of points. We may have a survey where we asked n people the same p questions. We can put that data in a matrix of dimensions $n \times p$, where each row is a person and each column is one of the asked questions.

Before we saw how to put a line through two points ($y = mx + b$). What if we wanted to put a line through many points?

Least Squares

Example



Least Squares

So how do we choose this line?

Least Squares

So how do we choose this line?
We can write the equation:

$$y \approx \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

where we have an intercept β_0 and then a slope β_i for each x_i (where $i = 1, .., p$). This equation has to describe the relationship as best it can for all n people we asked.

Least Squares

So how do we choose this line?
We can write the equation:

$$y \approx \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

where we have an intercept β_0 and then a slope β_i for each x_i (where $i = 1, .., p$). This equation has to describe the relationship as best it can for all n people we asked. In matrix notation:

Least Squares

So how do we choose this line?
We can write the equation:

$$y \approx \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

where we have an intercept β_0 and then a slope β_i for each x_i (where $i = 1, .., p$). This equation has to describe the relationship as best it can for all n people we asked. In matrix notation:

$$\begin{bmatrix} y_1 \\ ... \\ y_n \end{bmatrix} \approx \begin{bmatrix} 1 & x_{11} & ... & x_{1p} \\ 1 & ... & ... & ... \\ 1 & x_{n1} & ... & x_{np} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ ... \\ \beta_p \end{bmatrix}$$

or $y \approx X\beta$.

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$y \quad \approx \quad X\beta$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned} y &\approx X\beta \\ X^t y &\approx X^t X\beta \end{aligned}$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned} y &\approx X\beta \\ X^t y &\approx X^t X\beta \\ \left(X^t X\right)^{-1} X^t y &\approx \left(X^t X\right)^{-1} X^t X\beta \end{aligned}$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned}
y &\approx X\beta \\
X^t y &\approx X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx \left(X^t X\right)^{-1} X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx I \cdot \beta
\end{aligned}$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned}
y &\approx X\beta \\
X^t y &\approx X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx \left(X^t X\right)^{-1} X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx I \cdot \beta \\
\beta &\approx \left(X^t \cdot X\right)^{-1} X^t y
\end{aligned}$$