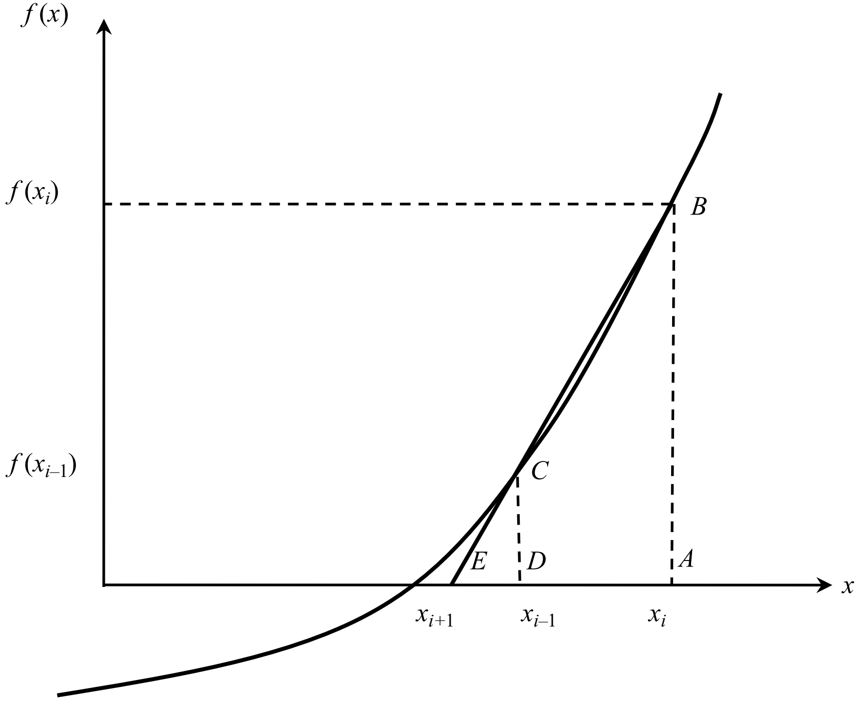
Secant Method

Secant Method

- Secant method is an open method. That is, no more interval under consideration needs to bracket the root, though it still requires two initial guesses. Let us denote these initial guesses by x_{-1} and x_0 . The formula for generating the sequence of approximations is



$$x_{i+1} = \frac{x_{i-1} f(x_i) - x_i f(x_{i-1})}{f(x_i) - f(x_{i-1})}$$

Secant Method

- Firstly, initial choices of $x_{-1}$and x_0 need not be bracketing the root. So, that is, root need not lie within the endpoints $x_{-1}$and x_0.That is f(x_{-1})and f(x_0)can be of same sign. No more, one needs to ensure that f(x_{-1}).f(x_0) < 0.

- Though, in practice one usually chooses $x_{-1}$and x_0 as ones which bracket the root, there is no such compulsion.

Secant Method

- Secondly, in computation of next approximation, older approximation is discarded. Recent most two approximations are used for iteration, irrespective of function values at these end points. So graphically, straight line joining (x_{i-1}, f (x_{i-1})) and (x_i , f (x_i)) is drawn to generate x_{i+1} , next approximation. No more, function values sign are checked to ensure that root lies between (x_{i-1}, f (x_{i-1})) and (x_i , f (x_i)) . Because secant is drawn joining (x_{i-1}, f (x_{i-1})) and (x_i , f (x_i)), it is called Secant Method.

Algorithm

The steps of the Secant method to find the root of an equation f(x) = 0 are

1. Let k =1, Evaluate $f(x_i)$, $f(x_{i-1})$
2. Use an initial guess of the root, x_i , to estimate the new value of the root, x_{i+1} , as

$$x_{i+1} = \frac{x_{i-1}f(x_i) - x_i f(x_{i-1})}{f(x_i) - f(x_{i-1})}$$

3. Compare the absolute approximate error $|c_k - c_{k-1}|$ or absolute relative approximate error $\frac{|c_k - c_{k-1}|}{|c_k|}$

with the pre-specified relative error tolerance ϵ.

where

c_k = estimated root from present iteration

c_{k-1}= estimated root from previous iteration

$If\ |c_k - c_{k-1}| < \epsilon\ or\ \frac{|c_k - c_{k-1}|}{|c_k|} < \epsilon$ then exit

Else

k = k + 1

go to Step 2

Note one should also check whether the number of iterations is more than the maximum number of iterations allowed. If so, one needs to terminate the algorithm and notify the user about it.

Advantages and Disadvantages

Advantages:

1. No constraint of end points of interval to converge
2. Very rapid convergence

Disadvantages:

1. If initial guess is not sufficiently near the root, it may diverge.

Example

This polynomial obviously has roots at x = 2 and at x = 4; one is a double root, the other a triple root:

$f(x) = (x-2)^3(x-4)^2$
$= x^5 - 14x^4 + 76x^3 - 200x^2 + 256x - 128$

a. Which root can you get with bisection? Which root can't you get?

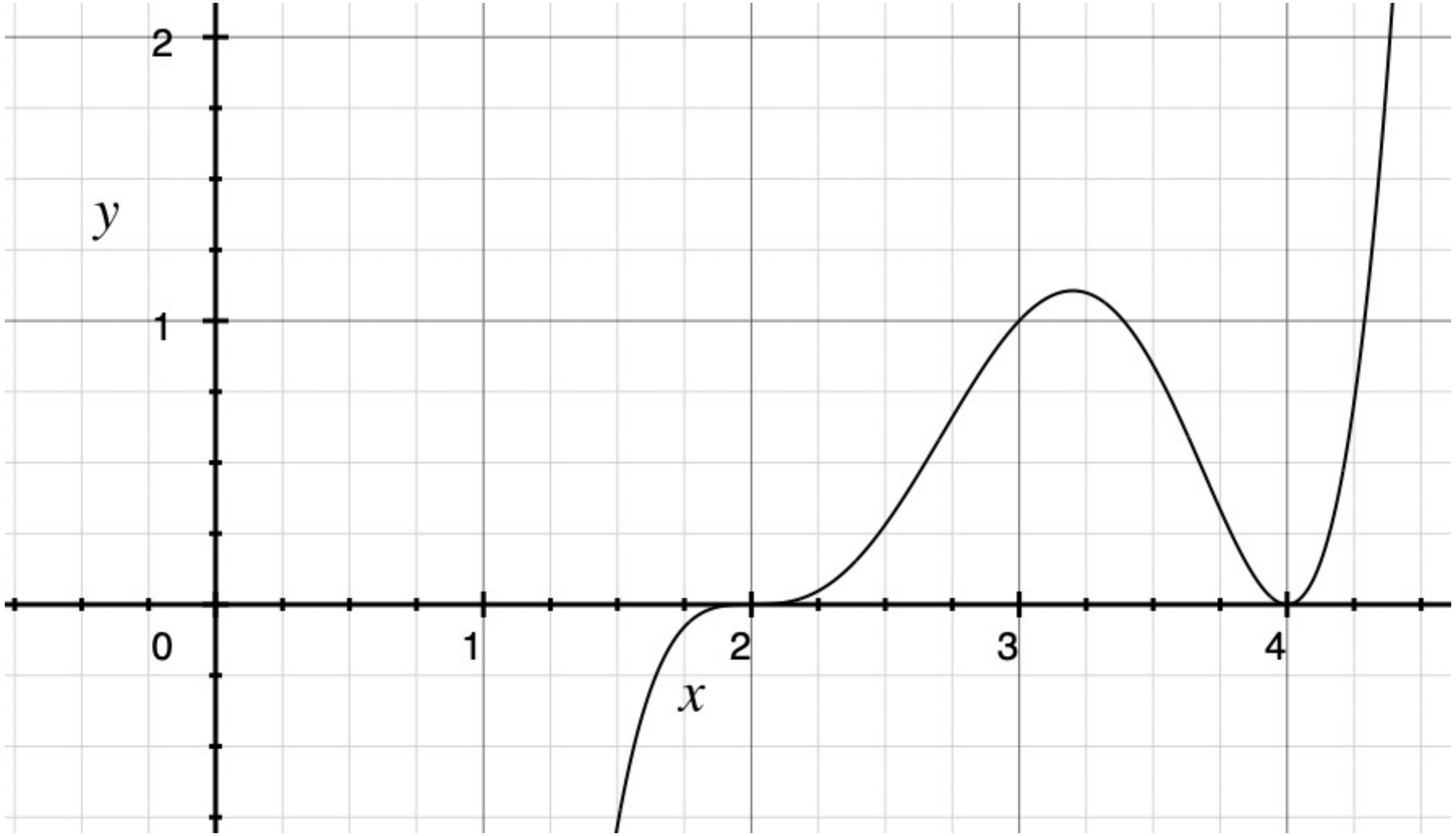
b. Repeat part (a) with the secant method.

c. If you begin with the interval [1,5], which root will you get with

(1) bisection, (2) the secant method, (3) false position?

d. Use Newton's method with x_0 = 3. Does it converge? To which root?

Example



Example

This polynomial obviously has roots at x = 2 and at x = 4; one is a double root, the other a triple root:

$f(x) = (x-2)^3(x-4)^2$

$= x^5 - 14x^4 + 76x^3 - 200x^2 + 256x - 128$

a.

b. Both roots can be reached with secant method

c. If you begin with the interval [1,5], which root will you get with

(2) the secant method: It will converge to 2

d.

Example

Eg 1. *Find the real root of the equation* $x^3 - 4x - 9 = 0$ *by* Secant Method *correct.*

Take x_0 *= 2,* x_1 *= 3,* $\epsilon = 0.01$

Eg 2. *Find the real root of the equation* $x^3 - x - 1 = 0$ *by* Secant Method *correct. Take* x_0 *= 1,* x_1 *= 1.5,* $\epsilon = 0.001$

Eg 3. $f(x) = 3x - e^x$ starting from 0,1

Secant Method

Eg 1. *Find the real root of the equation* $x^3 - 4x - 9 = 0$ *by* Secant Method *correct.*

x	0	1	2	3
f(x)	-9	-12	-9	6

Take x_0 = *2,* x_1 = *3,* ϵ = 0.01

Secant Method

Eg 1. *Find the real root of the equation $x^3 - 4x - 9 = 0$ by* Secant Method *correct. Take x_0 = 2, x_1 = 3,* ϵ = 0.001

$f(x) = x^3 - 4x - 9$

$f(2) = -9$

and

$f(3) = 6$

x_0	f(x_0)	x_1	f(x_1)
2	-9	3	6

Secant Method

First approximation to the root is

$$x_2 = \frac{(2)(6) - (3)(-9)}{6 - (-9)}$$

x_2= 2.6

Now $f(x_2)$ = – 1.824

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\lvert x_k - x_{k-1} \rvert$
2	-9	3	6	2.6	– 1.824	-

Secant Method

x_0 = *3*, x_1 = 2.6

Second approximation to the root is

$$x_2 = \frac{(3)(-1.824) - (2.6)6}{(-1.824) - 6}$$

= 2.6933

Now $f(x_2)$ = -0.2372

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\|x_k - x_{k-1}\|$
2	-9	3	6	2.6	– 1.824	-
3	6	2.6	– 1.824	2.6933	-0.2372	0.0933

Secant Method

x_0 = 2.6, x_1 = 2.6933

Third approximation to the root is

$$x_2 = \frac{(2.6)(-0.2372) - (2.6933)(-1.824\ )}{(-0.2372) - (2.6933)}$$

= 2.7072

Now $f(x_2)$ = 0.012

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\lvert x_k - x_{k-1} \rvert$
2	-9	3	6	2.6	– 1.824	-
3	6	2.6	– 1.824	2.6933	-0.2372	0.0933
2.6	– 1.824	2.6933	-0.2372	2.7072	0.012	0.0139

Secant Method

x_0 = 2.6933, x_1 = 2.7072

Fourth approximation to the root is

$x_2 = \frac{(2.6933)(0.012)-(2.7072)(-0.2372)}{0.012-(-0.2372)}$ = 2.7065

ϵ = 0.001, $|x_k - x_{k-1}|$ = 0.0007

Hence, the root is 2.7065, correct to three decimal places.

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\lvert x_k - x_{k-1} \rvert$
2	-9	3	6	2.6	– 1.824	-
3	6	2.6	– 1.824	2.6933	-0.2372	0.0933
2.6	– 1.824	2.6933	-0.2372	2.7072	0.012	0.0139
2.6933	-0.2372	2.7072	0.012	2.7065	0	0.0007

Secant Method

Eg 2. *Find the real root of the equation $x^3 - x - 1 = 0$ by* Secant Method *correct. Take* x_0 *= 1,* x_1 *= 1.5,* $\epsilon = 0.001$

$f(x) = x^3 - x - 1$

$f(1) = -1$

and

$f(1.5) = 0.875$

x_0	f(x_0)	x_1	f(x_1)
1	-1	1.5	0.875

Secant Method

First approximation to the root is

$$x_2 = \frac{(1)(0.875) - (1.5)(-1)}{0.875 - (-1)}$$

x_2= 1.2667

Now $f(x_2)$ = – 0.2344

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\|x_k - x_{k-1}\|$
1	-1	1.5	0.875	1.2667	– 0.2344	-

Secant Method

x_0 = *1.5*, x_1 = 1.2667

Second approximation to the root is

$$x_2 = \frac{(1.5)(-0.2344) - (1.2667)0.875}{(-0.2344) - 0.875}$$

= 1.316

Now $f(x_2)$ = -0.037

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\lvert x_k - x_{k-1} \rvert$
1	-1	1.5	0.875	1.2667	– 0.2344	-
1.5	0.875	1.2667	– 0.2344	1.316	-0.037	0.0493

Secant Method

x_0 = 1.2667, x_1 = 1.316

Third approximation to the root is

$$x_2 = \frac{(1.2667)(-0.037) - (1.316)(-0.2344)}{(-0.037) - (-0.2344)}$$

=1.3252

Now $f(x_2) = 0.0021$

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\lvert x_k - x_{k-1} \rvert$
1	-1	1.5	0.875	1.2667	– 0.2344	-
1.5	0.875	1.2667	– 0.2344	1.316	-0.037	0.0493
1.2667	– 0.2344	1.316	-0.037	1.3252	0.0021	0.0092

Secant Method

x_0 = 1.316, x_1 = 1.3252

Fourth approximation to the root is

$$x_2 = \frac{(1.316)(0.0021)-(1.3252)(-0.037)}{0.0021-(-0.037)} = 1.3247$$

ϵ = 0.001, $|x_k - x_{k-1}|$ = 0.0005

Hence, the root is 1.3247, correct to three decimal places.

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\lvert x_k - x_{k-1} \rvert$
1	-1	1.5	0.875	1.2667	– 0.2344	-
1.5	0.875	1.2667	– 0.2344	1.316	-0.037	0.0493
1.2667	– 0.2344	1.316	-0.037	1.3252	0.0021	0.0092
1.316	-0.037	1.3252	0.0021	1.3247	0	0.0005

Example

- f(x) = 3x - e^x starting from 0,1

x_0	f(x_0)	x_1	f(x_1)	x_2	f(x_2)	$\lvert x_k - x_{k-1} \rvert$
0	-1	1	0.2817	0.7802	0.1587	-
1	0.2817	0.7802	0.1587	0.4967	-0.1532	0.2835
0.7802	0.1587	0.4967	-0.1532	0.636	0.019	0.1393
0.4967	-0.1532	0.636	0.019	0.6206	0.0017	0.0154
0.636	0.019	0.6206	0.0017	0.619	0	0.0016

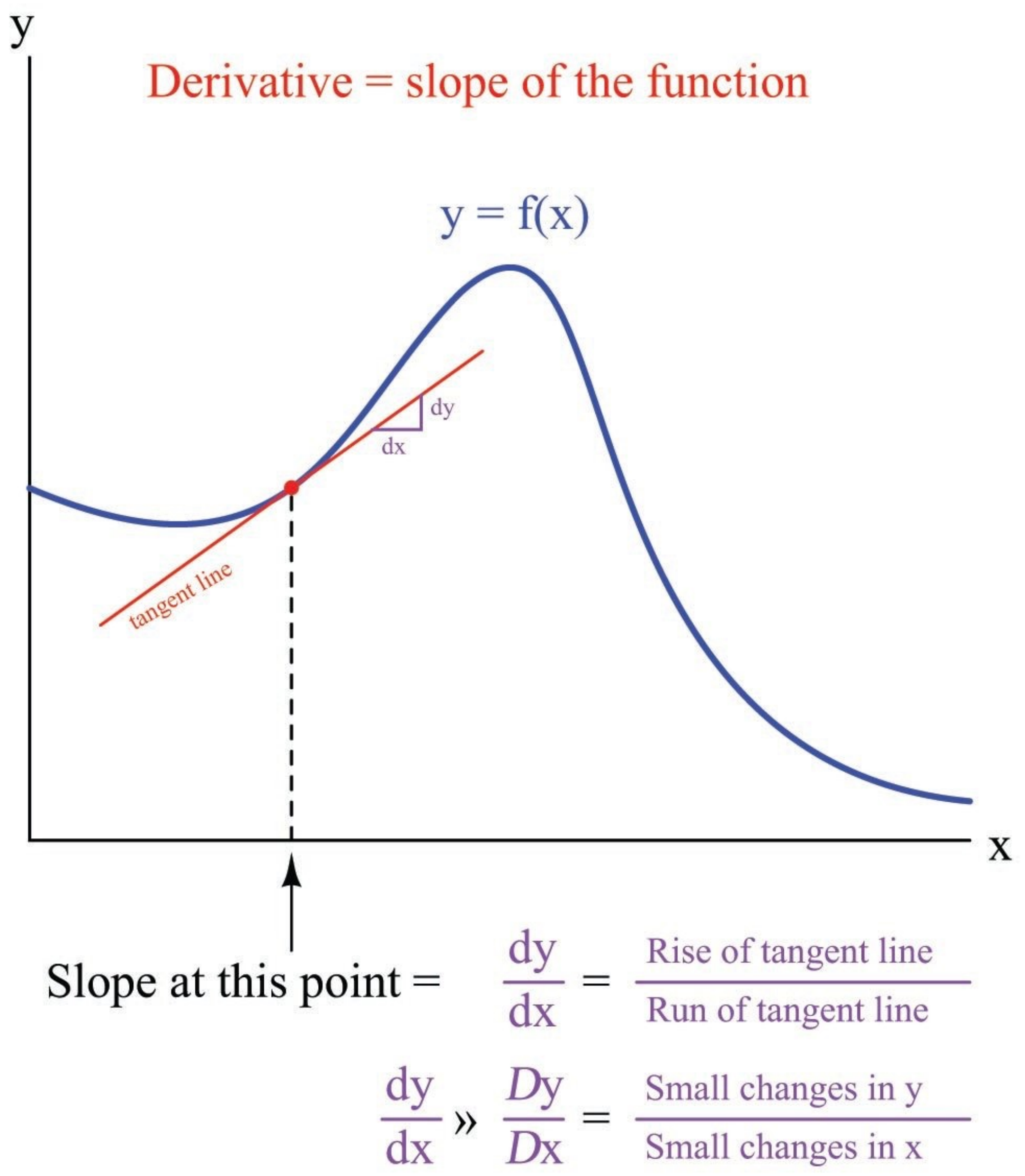
Gradient Descent

Derivative

- The derivative of a function at a chosen input value describes the rate of change of the function near that input value. The process of finding a derivative is called differentiation.

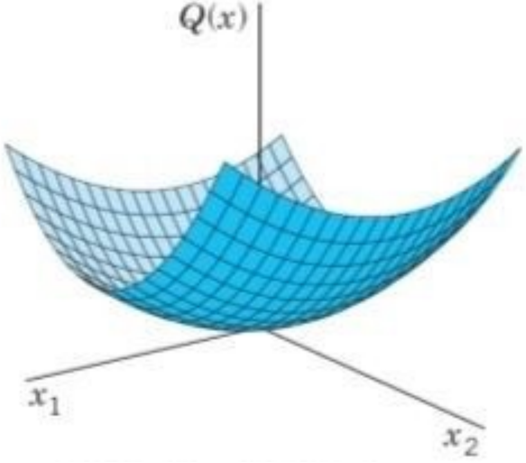
Derivative

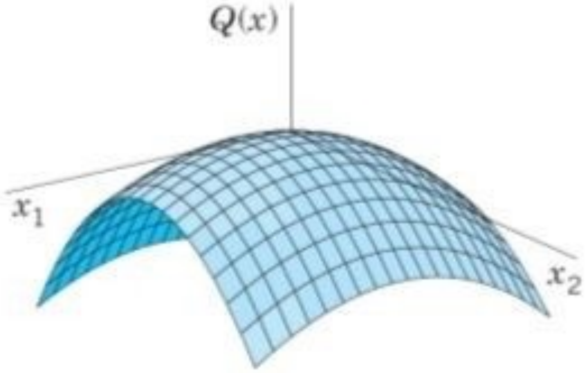
- Geometrically, the derivative at a point is the slope of the tangent line to the graph of the function at that point, provided that the derivative exists and is defined at that point. For a real valued function of a single real variable, the derivative of a function at a point generally determines the best linear approximation to the function at that point.



Gradient descent

- Optimization technique called **gradient descent**, which has seen major application in machine learning models.
- Gradient descent is an optimization technique that can find the *minimum* of an **objective function**. It is a greedy technique that finds the optimal solution by taking a step in the direction of the maximum rate of decrease of the function.





Gradient descent

- The basic intuition behind gradient descent can be illustrated by a hypothetical scenario. A blind-folded person is stuck in the mountains and is trying to get down (i.e., trying to find the global minimum). There is heavy fog such that visibility is extremely low. Therefore, the path down the mountain is not visible, so they must use local information to find the minimum.

Gradient descent

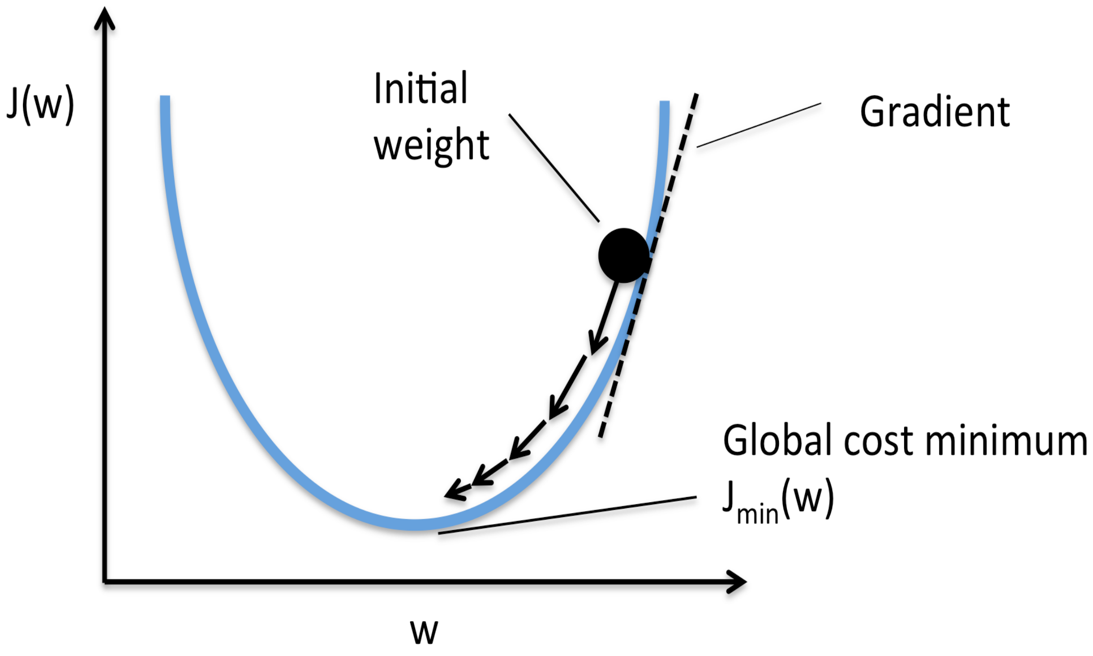
- They can use the method of gradient descent, which involves looking at the steepness of the hill at their current position, then proceeding in the direction with the steepest descent (i.e., downhill). If they were trying to find the top of the mountain (i.e., the maximum), then they would proceed in the direction of steepest ascent (i.e., uphill).

Gradient descent

- Using this method, they would eventually find their way down the mountain or possibly get stuck in some hole (i.e., local minimum or saddle point), like a mountain lake. However, assume also that the steepness of the hill is not immediately obvious with simple observation, but rather it requires a sophisticated instrument to measure, which the person happens to have at the moment. It takes quite some time to measure the steepness of the hill with the instrument, thus they should minimize their use of the instrument if they wanted to get down the mountain before sunset. The difficulty then is choosing the frequency at which they should measure the steepness of the hill so not to go off track.

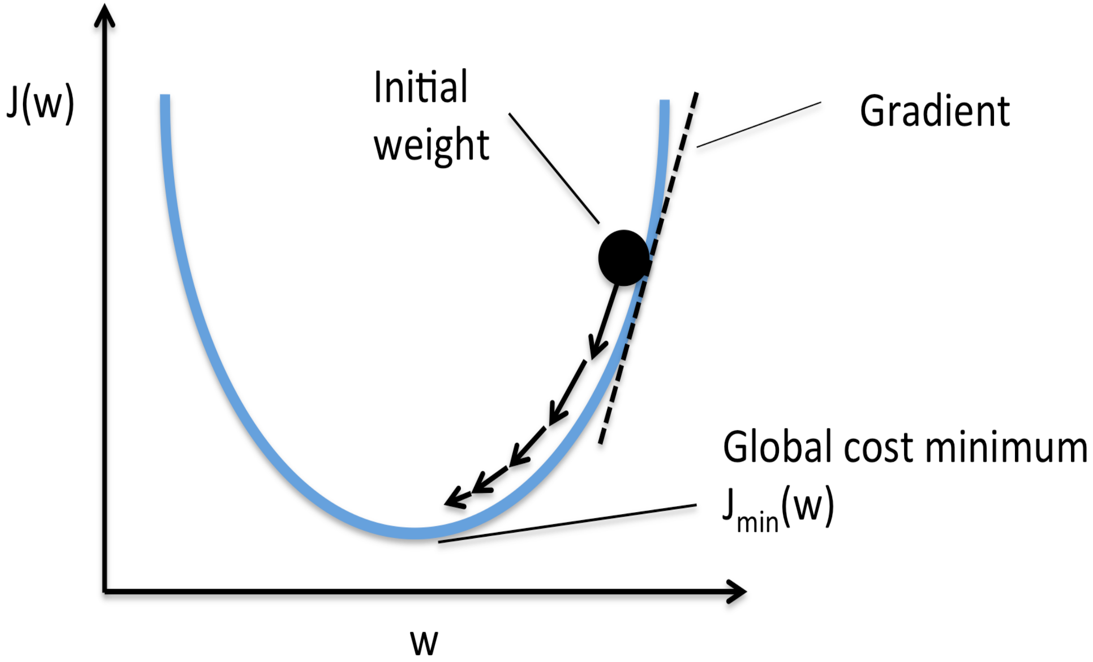
Gradient descent

- In this analogy, the person represents the algorithm, and the path taken down the mountain represents the sequence of parameter settings that the algorithm will explore. The steepness of the hill represents the <u>slope</u> of the error surface at that point. The instrument used to measure steepness is <u>differentiation</u>



Gradient descent

- The direction they choose to travel in aligns with the gradient of the error surface at that point. The amount of time they travel before taking another measurement is the step size.



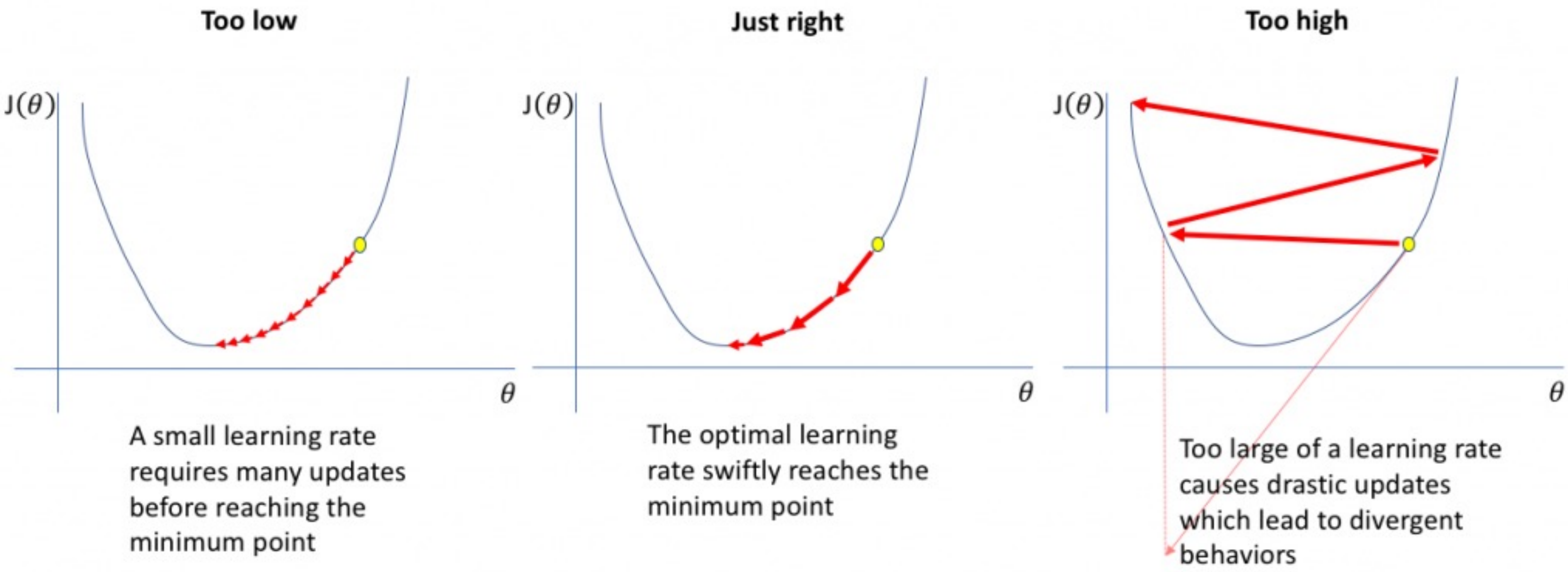
Gradient descent

- A gradient simply measures the change in all weights with regard to the change in error. You can also think of a gradient as the slope of a function. The higher the gradient, the steeper the slope and the faster a model can learn. But if the slope is zero, the model stops learning. In mathematical terms, a gradient is a partial derivative with respect to its inputs

Learning Rate

- Importance of the Learning Rate
- How big the steps are gradient descent takes into the direction of the local minimum are determined by the learning rate, which figures out how fast or slow we will move towards the optimal weights.
- For gradient descent to reach the local minimum we must set the learning rate to an appropriate value, which is neither too low nor too high. This is important because if the steps it takes are too big, it may not reach the local minimum because it bounces back and forth between the convex function of gradient descent. If we set the learning rate to a very small value, gradient descent will eventually reach the local minimum but that may take a while

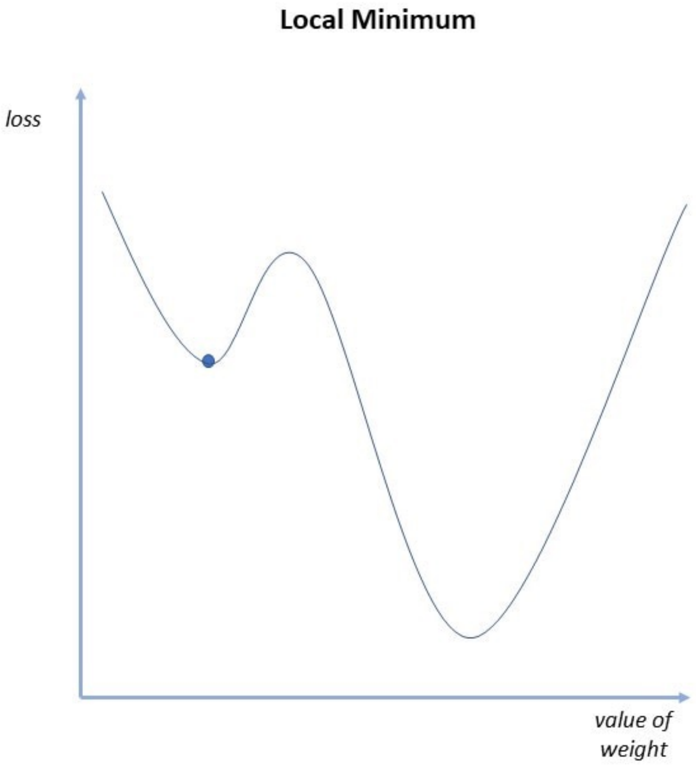
Learning Rate

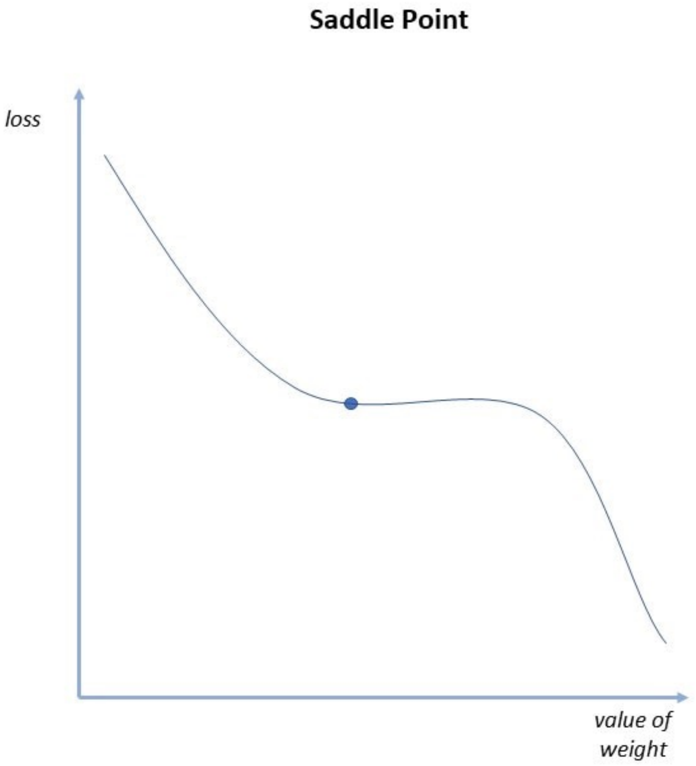


Local minima and saddle points

- **Challenges with gradient descent**
- Local minima and saddle points
- For convex problems, gradient descent can find the global minimum with ease, but as nonconvex problems emerge, gradient descent can struggle to find the global minimum, where the model achieves the best results.
- Recall that when the slope of the cost function is at or close to zero, the model stops learning. A few scenarios beyond the global minimum can also yield this slope, which are local minima and saddle points. Local minima mimic the shape of a global minimum, where the slope of the cost function increases on either side of the current point. However, with saddle points, the negative gradient only exists on one side of the point, reaching a local maximum on one side and a local minimum on the other. Its name inspired by that of a horse's saddle.
- Noisy gradients can help the gradient escape local minimums and saddle points.

Local minima and saddle points





Eigen vector and Eigen value

David Lay

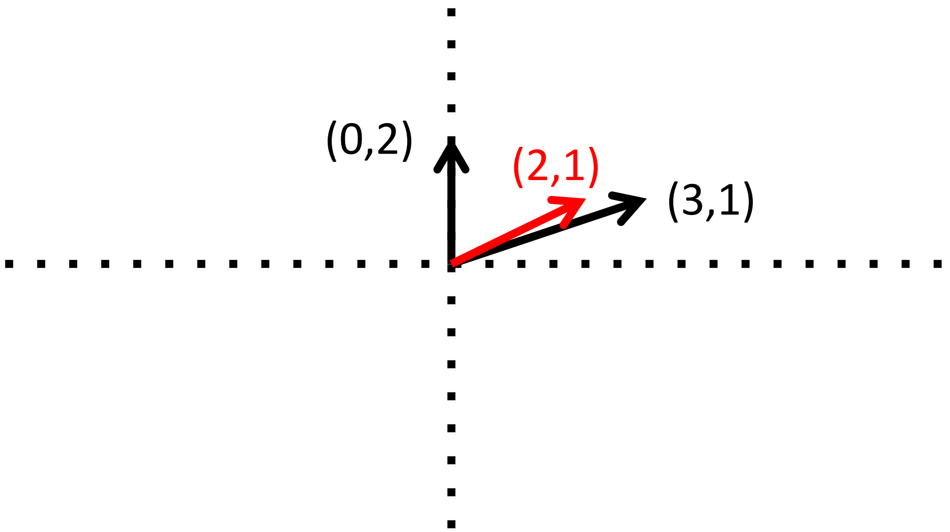
Eigen Value and Eigen Vectors

- An eigenvector of an $n \times n$ matrix A is a nonzero vector x such that $Ax = \lambda x$ for some scalar λ.
- A scalar λ is called an eigenvalue of A if there is a nontrivial solution x of $Ax = \lambda x$; such an x is called an eigenvector corresponding to λ.

What do matrices do to vectors?

$$\overleftrightarrow{M}$$

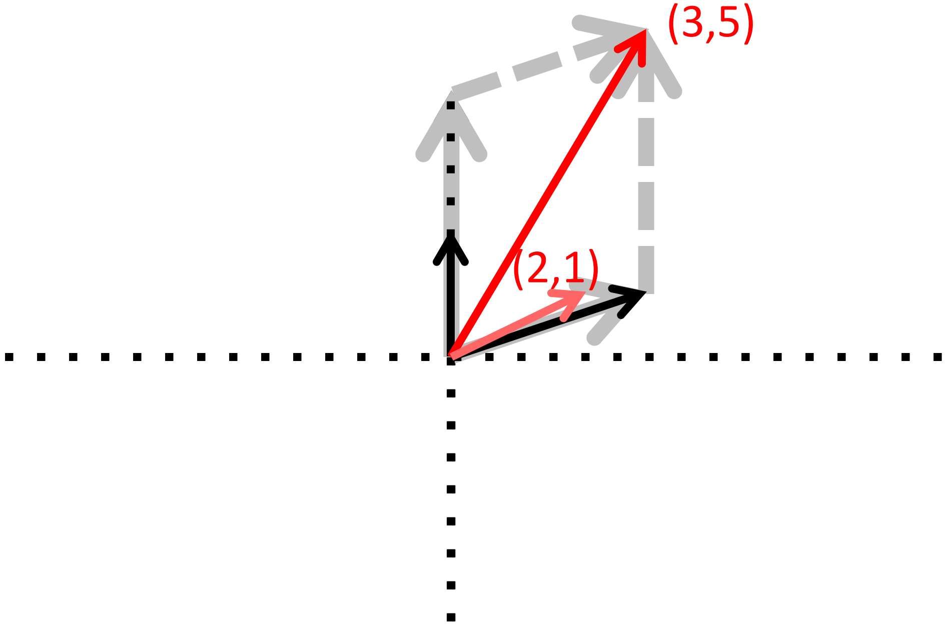
$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} =$$



What do matrices do to vectors?

$$\overleftrightarrow{M}$$

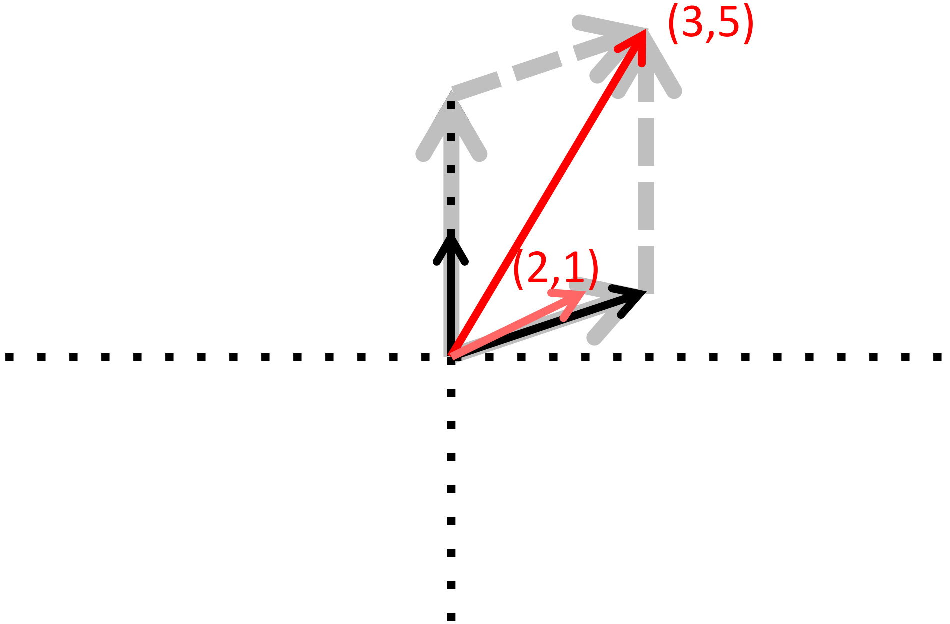
$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$



What do matrices do to vectors?

$$\overleftrightarrow{M}$$

$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

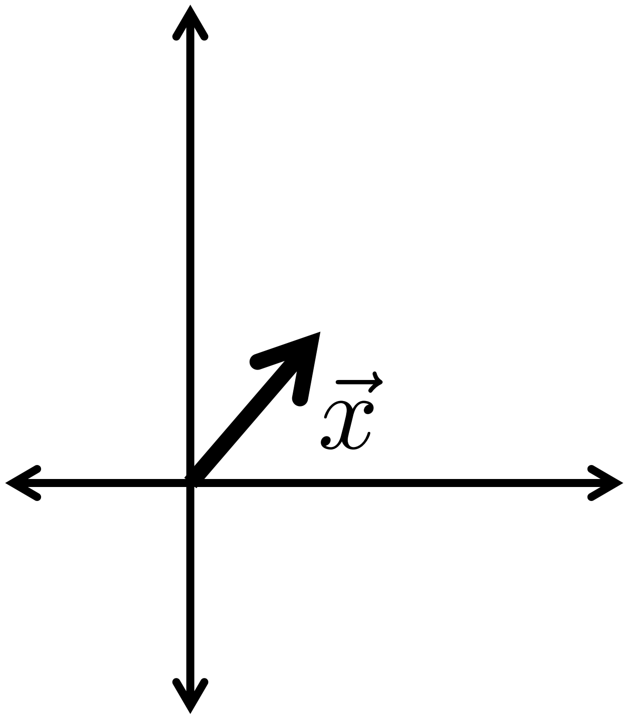


- The new vector is:
  1) **rotated**
  2) **scaled**

Are there any special vectors that **only get scaled**?

$\overleftrightarrow{M}$

$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

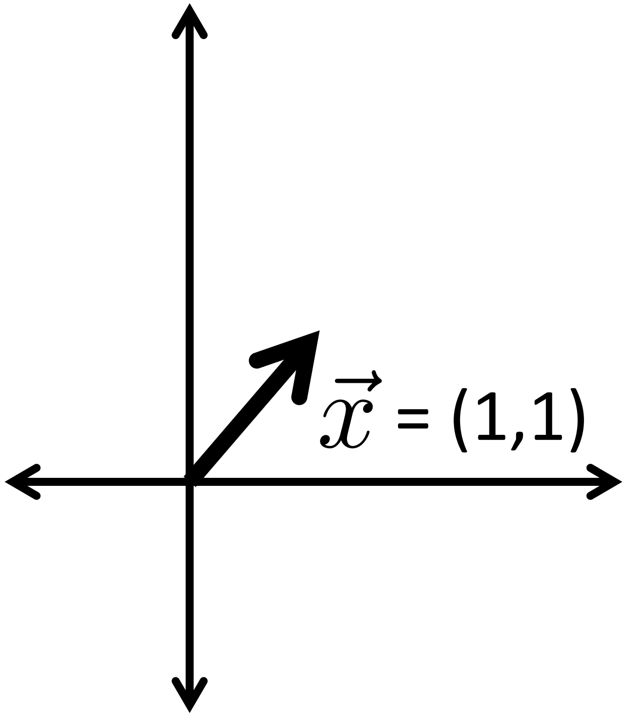


Try (1,1)

Are there any special vectors that **only get scaled**?

$\overleftrightarrow{M}$

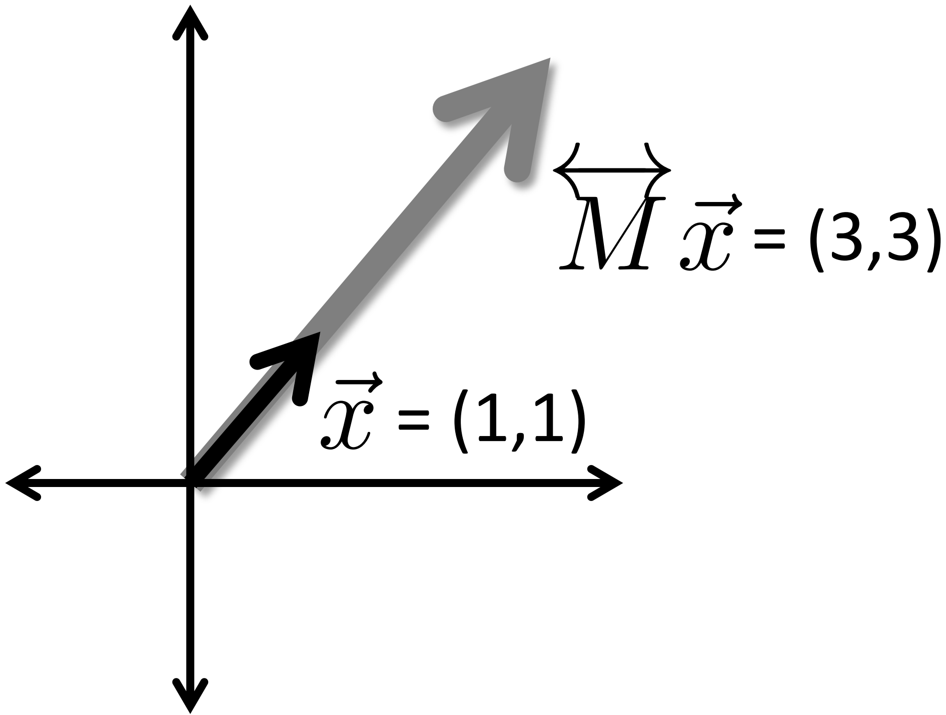
$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$



Are there any special vectors that **only get scaled**?

$\overleftrightarrow{M}$

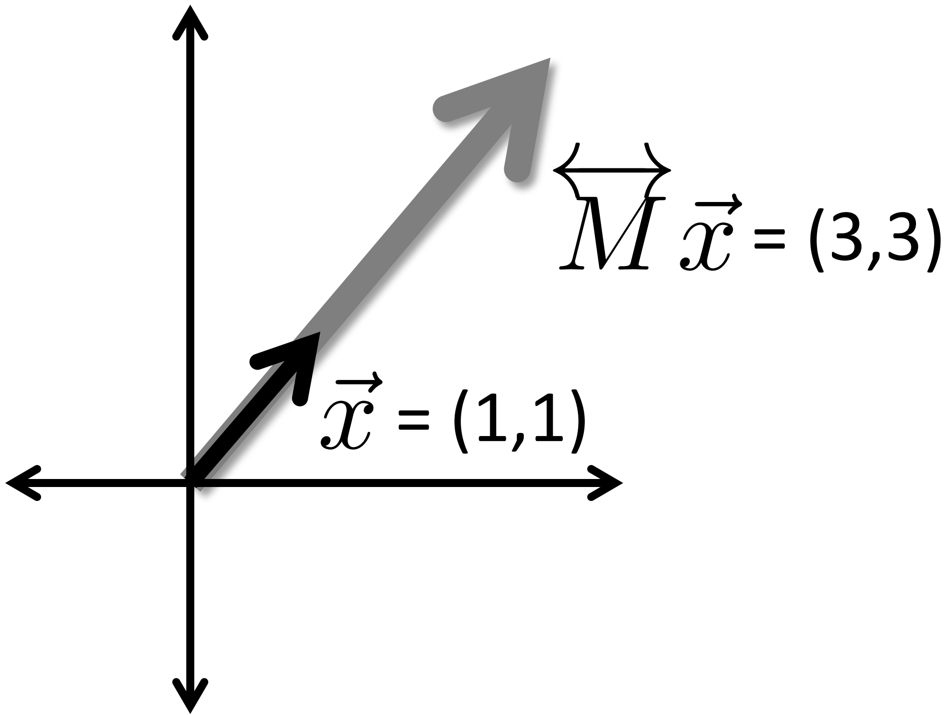
$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



Are there any special vectors that **only get scaled**?

$\overleftrightarrow{M}$

$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



- For this special vector, multiplying by M is like multiplying by a scalar.
- (1,1) is called an **eigenvector** of M
- 3 (the scaling factor) is called the **eigenvalue** associated with this eigenvector

Eigenvalues and Eigenvectors

- Eigenvalue and eigenvector:

A : an $n \times n$ matrix

λ : a scalar

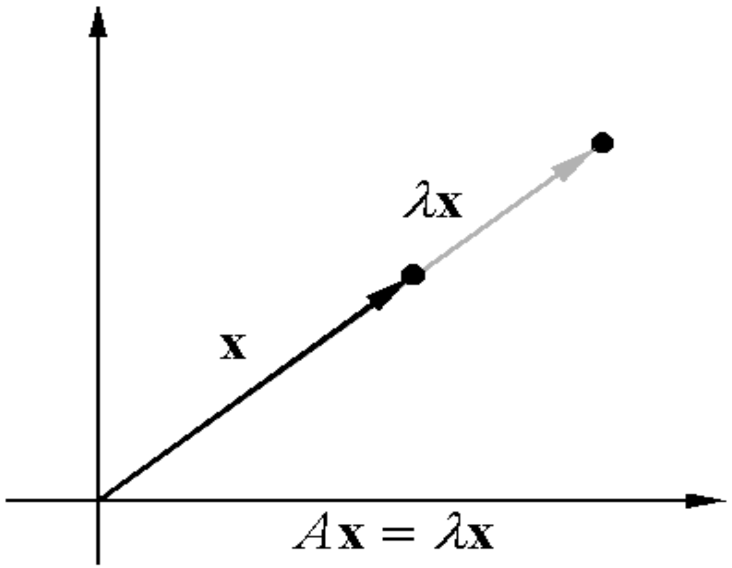
$\boldsymbol{x}$: <u>a nonzero vector</u> in R^n

Eigenvalue

$$Ax = \lambda x$$

Eigenvector

- Geometrical Interpretation



Example 1

Let A= $\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, u = $\begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and v = $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

Are u and v eigenvectors of A?

$Au =$ $\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} =$ -4$\begin{bmatrix} 6 \\ -5 \end{bmatrix}$ = -4u

Thus u is an eigenvector corresponding to an eigenvalue (−4)

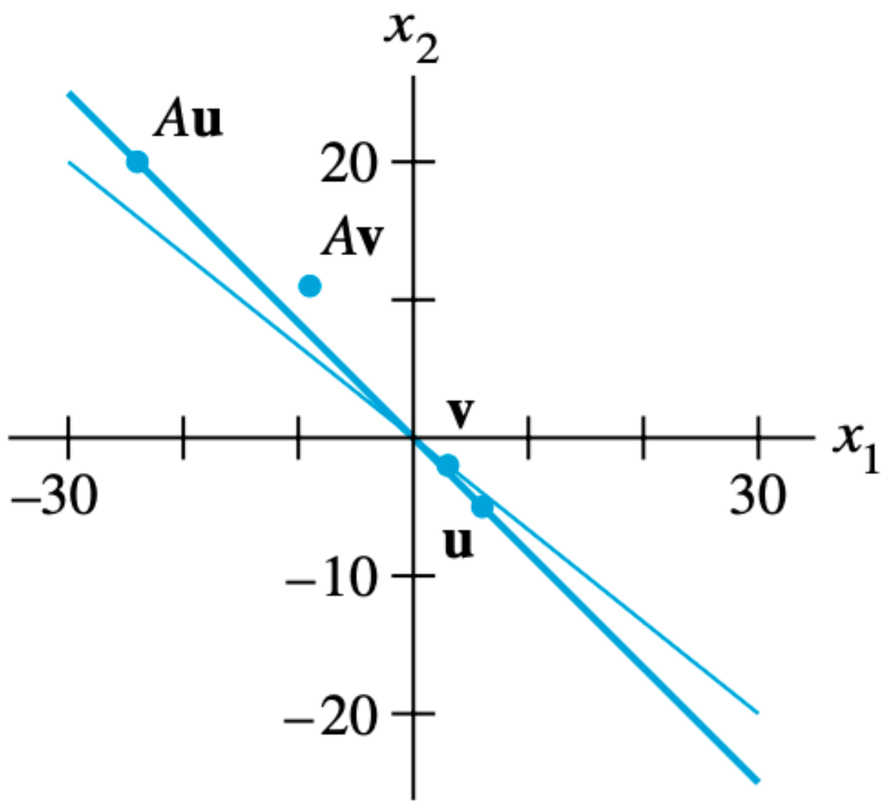
Example 1

Let A= $\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, u = $\begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and v = $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

Are u and v eigenvectors of A?

- $Av = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda\begin{bmatrix} 3 \\ -2 \end{bmatrix}$
- v is not an eigenvector of A, because Av is not a multiple of v.

Example 1



$A\mathbf{u} = -4\mathbf{u}$, but $A\mathbf{v} \neq \lambda\mathbf{v}$.

Example 2

Is $\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$ an eigen vector of $\begin{bmatrix} 3 & 7 & 9 \\ -4 & -5 & 1 \\ 2 & 4 & 4 \end{bmatrix}$

If so, find the eigenvalue.

Example 2

Is $\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$ an eigen vector of $\begin{bmatrix} 3 & 7 & 9 \\ -4 & -5 & 1 \\ 2 & 4 & 4 \end{bmatrix}$

If so, find the eigenvalue.

$$\begin{bmatrix} 3 & 7 & 9 \\ -4 & -5 & 1 \\ 2 & 4 & 4 \end{bmatrix}\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$ So $\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$ is an eigenvector of A

Eigen value is 0

Example 3

- Show that x = $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ and find the corresponding eigenvalue.

Example 3

Show that x = $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ and find the corresponding eigenvalue.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4\mathrm{x}$$

from which it follows that x is an eigenvector of A corresponding to the eigenvalue 4.

Eigen Space

λ is an eigenvalue of A if and only if the equation

$(A - \lambda I)x = 0$

has a nontrivial solution. The set of all solutions of the above equation is just the null space of the matrix $A - \lambda I$. So this set is a subspace of R^n and is called the eigenspace of A corresponding to λ. The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ.

All vectors in the nullspace of $A - \lambda I$ (which we call the eigenspace) will satisfy $Ax = \lambda x$.

Example 1

Let A = $\begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigen value of A is 2. Find a basis for the corresponding eigenspace.

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

Example 1

row reduce the augmented matrix for $(A - 2I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

2 is indeed an eigenvalue of A because the equation $(A - 2I)\mathbf{x} = \mathbf{0}$ has free variables. The general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \qquad x_2 \text{ and } x_3 \text{ free}$$

Example 1

The eigenspace, shown in Fig. 3, is a two-dimensional subspace of R3. A basis is

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

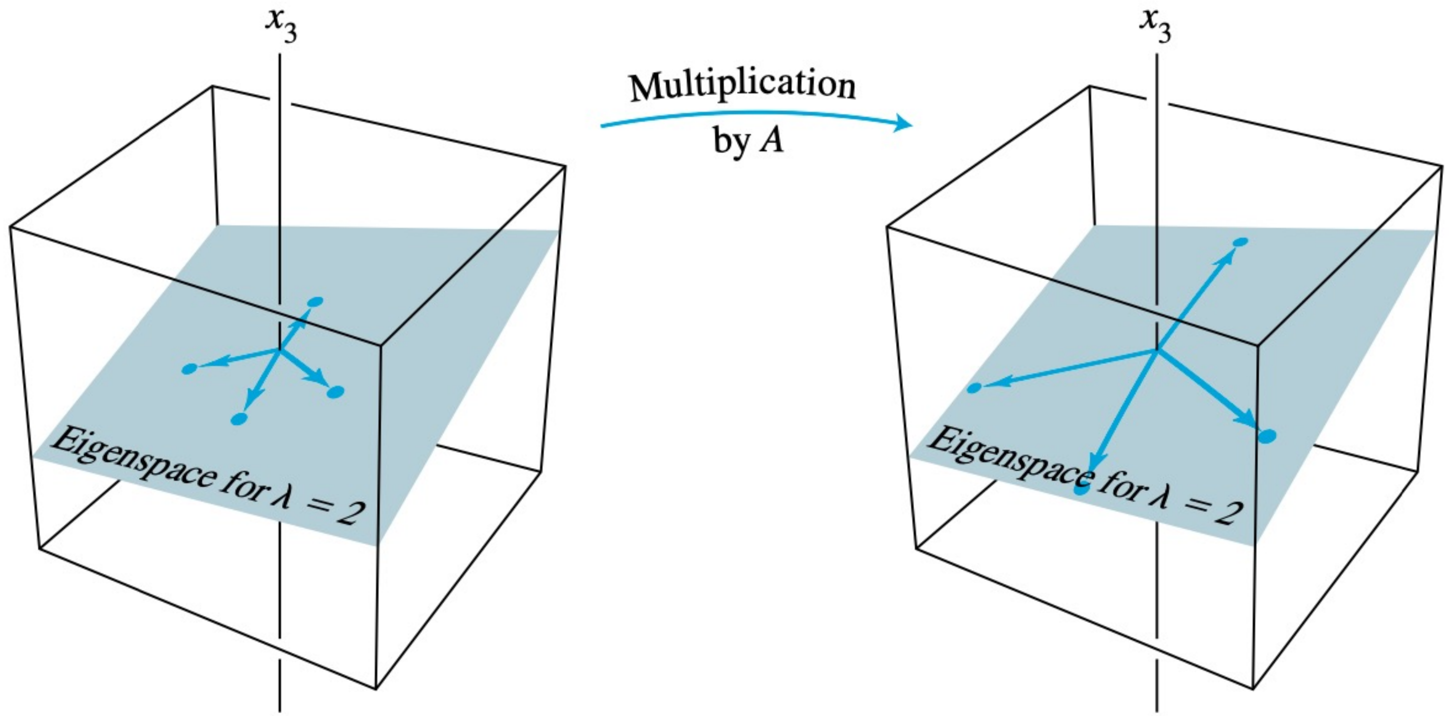


FIGURE 3 *A* acts as a dilation on the eigenspace.

Example 2

Show that 5 is an eigenvalue of $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ and determine all eigenvectors corresponding to this eigenvalue.

We must show that there is a nonzero vector x such that Ax = 5x. But this equation is equivalent to the equation (A - 5I)x = 0, so we need to compute the null space of the matrix A - 5I. We find that

$$A - 5I = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} -4 & 2 \\ 4 & -2 \end{bmatrix}$$

Since the columns of this matrix are clearly linearly dependent, its null space is nonzero.

Thus, Ax = 5x. has a nontrivial solution, so 5 is an eigenvalue of A.

Example 2

We find its eigenvectors by computing the null space:

$$[A - 5I \quad 0] = \begin{bmatrix} -4 & 2 & 0 \\ 4 & -2 & 0 \end{bmatrix} = \begin{bmatrix} -1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, if $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue 5, it satisfies

$-x_1 + \ 1/2x_2 = 0$ or $x_1 = \ 1/2x_2$

The eigen vector is $x_2 \begin{bmatrix} 1 \\ {}^1/_2 \end{bmatrix} = \begin{bmatrix} 1 \\ {}^1/_2 \end{bmatrix}$

The eigen space corresponding to λ = 5 consists of all multiples of $\begin{bmatrix} 1 \\ {}^1/_2 \end{bmatrix}$, which is the line through (1,1/2) and the origin.

Example 3

Is $\lambda = -2$ an eigenvalue of $\begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$? Why or why not?

Example 3

Is $\lambda = -2$ an eigenvalue of $\begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$? Why or why not?

The number -2 is an eigenvalue of A if and only if the equation $A\mathbf{x} = -2\mathbf{x}$ has a nontrivial solution. This equation is equivalent to $(A+2I)\mathbf{x} = \mathbf{0}$. Compute

$$A + 2I = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}$$

The columns of A are obviously linearly dependent, so $(A+2I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution, and so -2 is an eigenvalue of A.