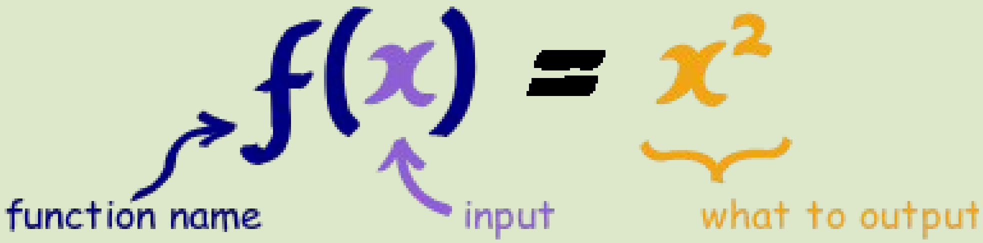
WQD7011
Numerical Optimization

Linear Algebra – a revision

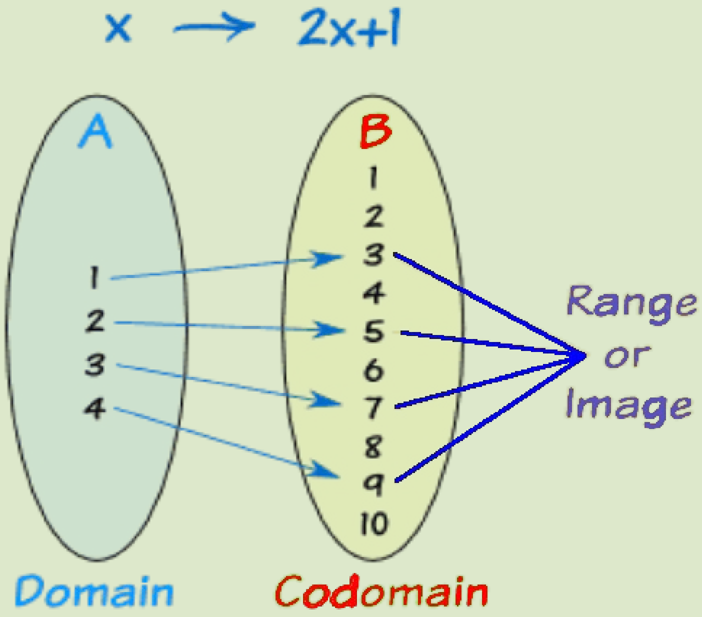
Content

- Functions
- Matrices
- Derivatives

Function



$f(x) = x^2$
function name
input
what to output



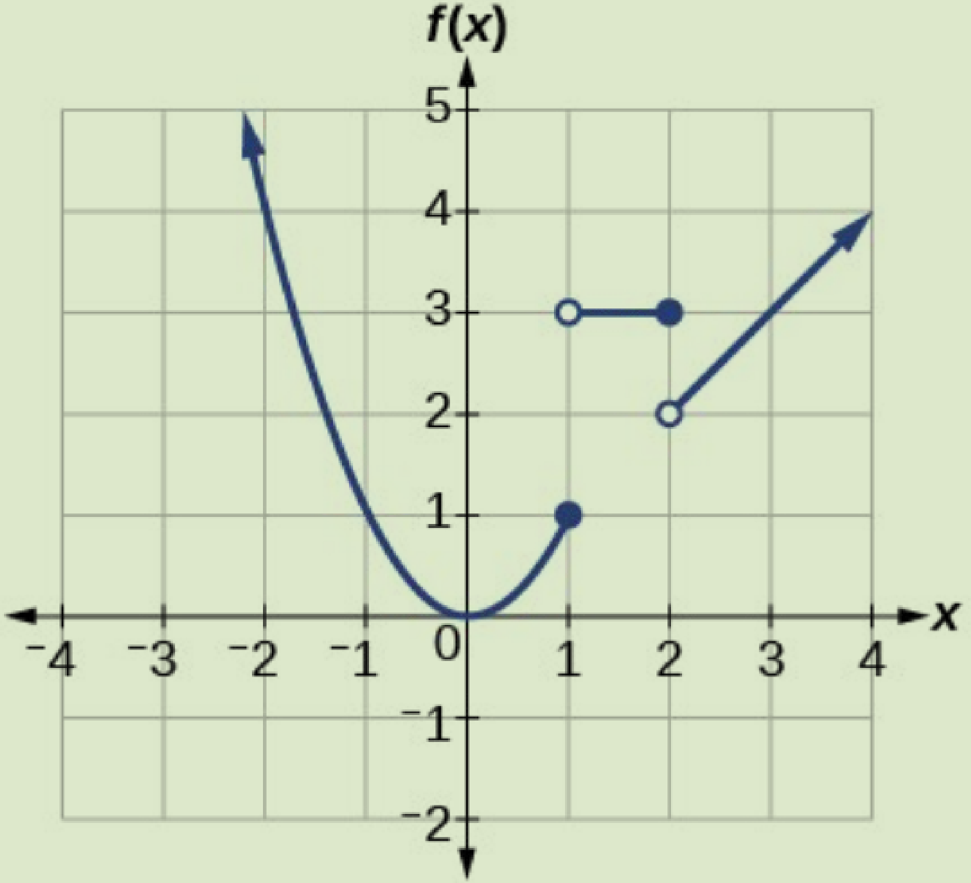
x → 2x+1
A
1
2
3
4
B
1
2
3
4
5
6
7
8
9
10
Range or Image
Domain
Codomain

Piecewise Functions

A function that behaves differently based on input x is called a piecewise function.

Example:

$$f(x) = \begin{cases} x^2, & if\ x \leq 1 \\ 3, & if\ 1 < x \leq 2 \\ x, & if\ x > 2 \end{cases}$$



Multivariable Functions

- Some functions have more than one variables.

  f: X x Y $\rightarrow$ Z

  where X and Y are domains and Z is codomain
- Sometimes, we can consider the variables of these as vectors.
- Examples:

  f(x,y) = x-y

  f(4,2) = 2

  f(3,-5) = 8

$$f(x,y) = \begin{cases} \frac{y}{x} - y & \text{if} \quad 0 \leq y < x \leq 1 \\ \frac{x}{y} - x & \text{if} \quad 0 \leq x < y \leq 1 \\ 1 - x & \text{if} \quad 0 < x = y \\ 0 & \text{everywhere else.} \end{cases}$$

Matrix

- A rectangular array of numbers.
- Example:

$$\begin{bmatrix} 3 & 4 \\ 7 & -2 \\ -5 & 0 \end{bmatrix}$$

Size of matrix : Row X Column

Matrix

- Generally, a matrix is in the form:

$$
\begin{array}{cc}
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\
a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\
\vdots & \vdots & \vdots & & \vdots \\
a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}
\end{bmatrix} & \left.\vphantom{\begin{bmatrix} a \\ a \\ a \\ \vdots \\ a \end{bmatrix}}\right\} m \text{ rows} \\
\underbrace{\hphantom{a_{11} \quad a_{12} \quad a_{13} \quad \cdots \quad a_{1n}}}_{n \text{ columns}} &
\end{array}
$$

Each entry, a_{ij} refers to the value in row i, column j.

If all entries for A are real numbers, $A \in \mathbb{R}^{mxn}$

Matrix Addition and Subtraction

We can add/subtract two matrices (of the same dimension) by adding their corresponding entries.

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

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

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

Scalar Multiplication

- We can multiply a matrix A by a scalar c by multiplying each entry in A by c

$$\underset{\text{Scalar}}{3}\underset{\text{Matrix}}{\begin{bmatrix} -1 & 2 \\ 6 & 5 \end{bmatrix}} = \begin{bmatrix} 3(-1) & 3(2) \\ 3(6) & 3(5) \end{bmatrix} = \begin{bmatrix} -3 & 6 \\ 18 & 15 \end{bmatrix}$$

Matrices Multiplication

- A matrices multiplication is a product of 2 matrices.

Requirement for AB:	Result for AB:
size of A is m x n and size of B is n x p	m x p matrix

If $A = [a_{ij}]$ and $B = [b_{ij}]$

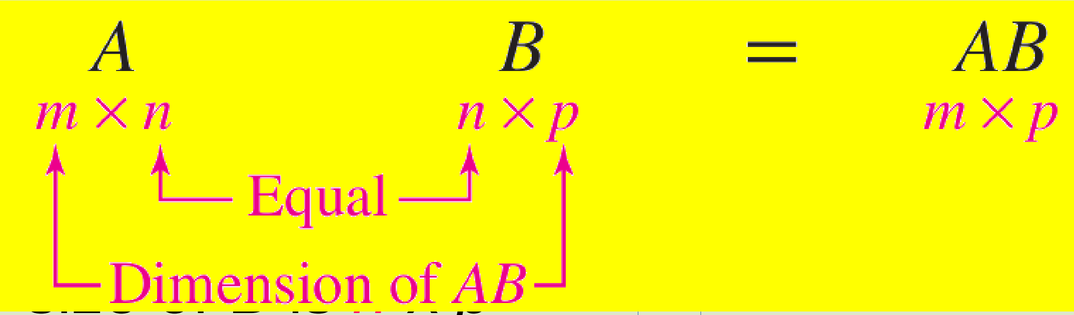
$AB = [c_{ij}]$ where $c_{ij} = a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + a_{i3}\, b_{3j} + \ldots + a_{in}\, b_{nj}$

Matrices Multiplication

- A matrices multiplication is a product of 2 matrices.

Re

si



AB:

atrix

If $A = [a_{ij}]$ and $B = [b_{ij}]$

$AB = [c_{ij}]$ where $c_{ij} = a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + a_{i3}\, b_{3j} + \ldots + a_{in}\, b_{nj}$

Matrices Multiplication

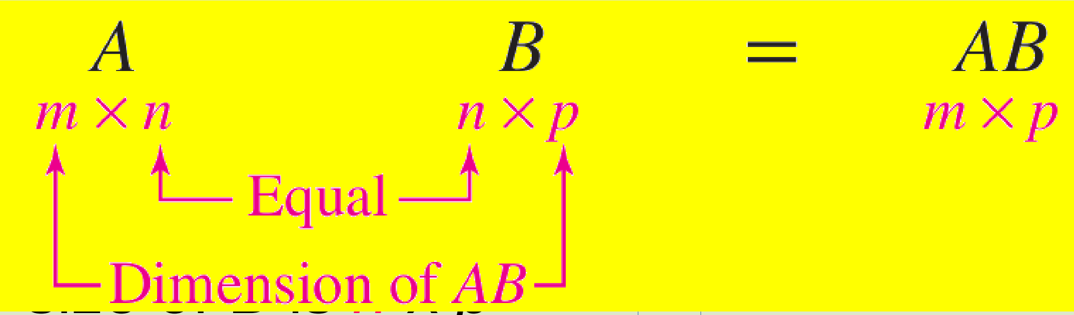
- A matrices multiplication is a product of 2 matrices.

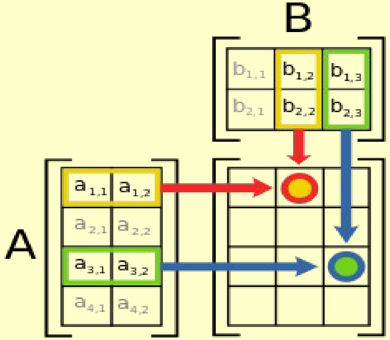
Re

si

: AB:

atrix





Square Matrix

- a square matrix is a n x n matrix where n is a positive Integer.
- An Identity matrix is a special case of square matrix, I_n.

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Diagonal Matrix

- All the entries not at the diagonal are 0.
- Identity matrix is a special case of diagonal matrix

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \qquad \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \end{bmatrix}$$

Symmetric Matrix

- Must be a square matrix,
- With symmetric property

$$\begin{bmatrix} 3 & 7 & -8 \\ 7 & -1 & 2 \\ -8 & 2 & 4 \end{bmatrix}$$

Upper / Lower Triangular Matrix

- Upper triangular matrix, U is defined as:

$$U_{ij} = \begin{cases} a_{ij} & \text{for } i \leq j \\ 0 & \text{for } i > j \end{cases}$$

$$\begin{bmatrix} 3 & 7 & -8 & 8 \\ 0 & -1 & 6 & 2 \\ 0 & 0 & 4 & 12 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

- Lower triangular matrix, U is defined as:

$$U_{ij} = \begin{cases} 0 & \text{for } i < j \\ a_{ij} & \text{for } i \geq j \end{cases}$$

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ -6 & -4 & 4 & 0 \\ 3 & -5 & 9 & 5 \end{bmatrix}$$

Transpose of Matrix

- If A is a m x n matrix, transpose of A, or A^T is a n x m matrix obtained from A by interchanging the rows and columns of A.
- After transpose,
  - Column i of A $\rightarrow$ Row i of A^T
  - Row j of A $\rightarrow$ Column j of A^T

$AB \neq BA$

$(AB)^T = B^T A^T$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Column and Row Vector

- When row =1, we have row vector

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

- When column =1, we have column vector

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

Column and Row Vector

- When row =1, we have row vector

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

- When column =1, we have column vector

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

Note: if $x \in \mathbb{R}^n$ $\quad A \in \mathbb{R}^{n \times n}$

$x^T x \in \mathbb{R}$ $\quad xx^T \in \mathbb{R}^{n \times n}$

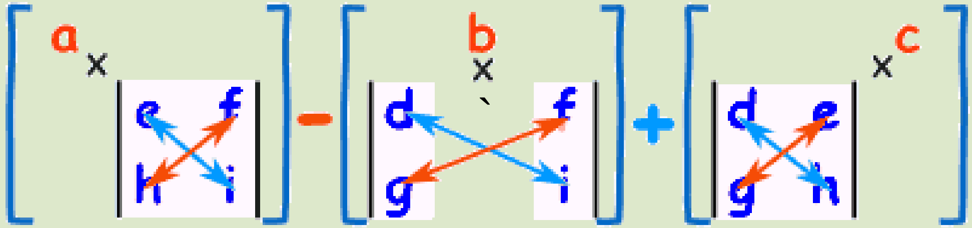
$x^T A x \in \mathbb{R}$

Determinant of a matrix

- If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, det(A) = ad – bc

Determinant of a matrix

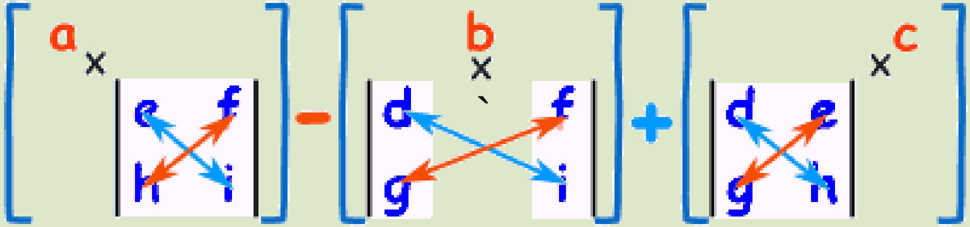
- If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, det(A) = ad – bc
- If $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$,



$$\det(A) = a(ei - fh) - b(di - fg) + c(dh - eg)$$

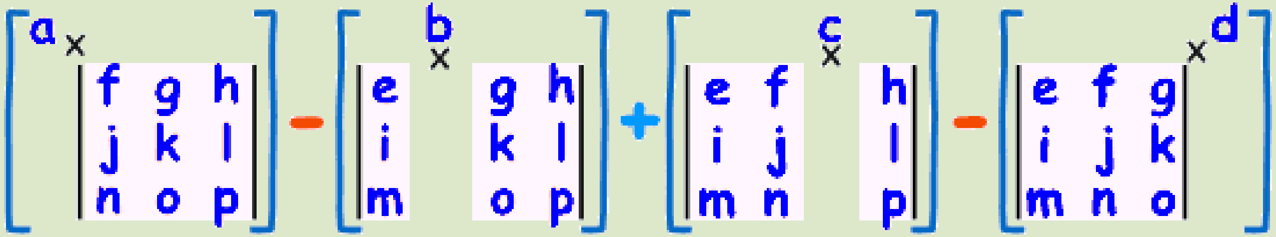
Determinant of a matrix

- If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, det(A) = ad – bc
- If $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$,



$$\det(A) = a(ei - fh) - b(di - fg) + c(dh - eg)$$

4X4 matrices:



Inverse of a matrix

- Let A and B are square matrices with dimensions n and In is the identity matrix.
- If $AB = BA = I_n$

  then B is the inverse of A.
- We can write B as A^{-1}
- $(A^{-1})^{-1}=A$
- $(AB)^{-1}=B^{-1}A^{-1}$

Inverse of a matrix

- Let A and B are square matrices with dimensions n and In is the identity matrix.
- If $AB = BA = I_n$

  then B is the inverse of A.
- We can write B as A^{-1}
- $(A^{-1})^{-1}=A$
- $(AB)^{-1}=B^{-1}A^{-1}$

If the determinant of a matrix is 0, the matrix has no inverse.

A matrix with no inverse is a singular matrix. Otherwise, it is nonsingular.

Positive Definite Matrix

- Given A $\in \mathbb{R}^{nxn}$, and x$\in \mathbb{R}^{n}$.
- A is positive definite if :

$$x^{\mathsf{T}}Ax > 0$$

- A is positive semidefinite if :

$$x^{\mathsf{T}}Ax \geq 0$$

- non symmetric matrices can be positive definite or semidefinite as well, but we don’t cover that

Positive Definite Matrix

- Symmetric *n*x*n* matrix A is positive definite if any of the following is fulfilled:
  - all its eigenvalues are positive.
  - For all $m \in [1, n]$, $\det(A^{(m)}) > 0$

    $A^{(m)}$ is the $m \times m$ matrix obtained by omitting all rows and columns of A past the *m*th.
  - Can be written as $A = R^T R$ (R can be rectangular)

Positive Definite Matrix (example)

- Assume $A = \begin{bmatrix} 7 & -4 & 1 \\ -4 & 5 & 0 \\ 1 & 0 & 6 \end{bmatrix}$

Eigenvalues of A =
{ 1.7854, 5.9444,
10.2702 }

All positive,
$\rightarrow$positive definite

Positive Definite Matrix (example)

- Assume $A = \begin{bmatrix} 7 & -4 & 1 \\ -4 & 5 & 0 \\ 1 & 0 & 6 \end{bmatrix}$

Eigenvalues of A = { 1.7854, 5.9444, 10.2702 }

All positive,
→positive definite

$$\det[7] = 7 > 0$$

$$\det \begin{bmatrix} 7 & -4 \\ -4 & 5 \end{bmatrix} = 19 > 0$$

$$\det \begin{bmatrix} 7 & -4 & 1 \\ -4 & 5 & 0 \\ 1 & 0 & 6 \end{bmatrix} = 109 > 0$$

All positive,
→ positive definite

Eigenvector and Eigenvalue

- When a square matrix *A* multiply with a nonzero vector *x*, it changes the direction of *x*, except when *x* is the eigenvector of *A*.

$$Ax = \lambda x$$

A : *n*x*n* matrix

x : eigenvector (non zero)

λ : eigenvalue

Eigenvector and Eigenvalue

$$Ax = \lambda x$$

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

Since x is non zero, det(A-λI) = 0

With this characteristic equation, λ can be found.

With the eigenvalues, eigenvectors can be found.

Eigenvector and Eigenvalue (example)

Given

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$$

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

$$\det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = 0$$

$$(2-\lambda)(-6-\lambda) - (3)(3) = 0$$

$$\lambda = 3, -7$$

Eigenvector and Eigenvalue (example)

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

$$\begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

For λ = 3,

$$\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$-x_1 + 3x_2 = 0$$

$$3x_1 - 9x_2 = 0$$

Both give $x_1 = 3x_2$

So, eigenvector corresponding to λ = 3 is:

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

With λ = -7, the eigenvector is $[\ 1\ 3\ ]^T$

Eigenvector and Eigenvalue

- The determinant of a matrix is the multiple of its eigenvalues.
- If a matrix has eigenvalue 0,it is a singular matrix.
- eigenvalues of symmetric matrices are all real numbers.
- nonsymmetric matrices may have imaginary eigenvalues.
- If the matrix is positive definite as well as symmetric, its eigenvalues are all positive real numbers.

Derivatives

- Let Φ : $\mathbb{R} \rightarrow \mathbb{R}$ be a real-valued function of a real variable.
- The first derivative $\Phi(\alpha)$ is defined by:

$$\frac{d\phi}{d\alpha} = \phi'(\alpha) \stackrel{\text{def}}{=} \lim_{\epsilon \to 0} \frac{\phi(\alpha + \epsilon) - \phi(\alpha)}{\epsilon}$$

- The second derivative is obtained by substituting Φ by Φ' in this same formula:

$$\frac{d^2\phi}{d\alpha^2} = \phi''(\alpha) \stackrel{\text{def}}{=} \lim_{\epsilon \to 0} \frac{\phi'(\alpha + \epsilon) - \phi'(\alpha)}{\epsilon}$$

Derivatives

Basic Properties and Formulas

If $f(x)$ and $g(x)$ are differentiable functions (the derivative exists), c and n are any real numbers,

1. $(cf)' = c f'(x)$
2. $(f \pm g)' = f'(x) \pm g'(x)$
3. $(fg)' = f'g + fg'$ – **Product Rule**
4. $\left(\dfrac{f}{g}\right)' = \dfrac{f'g - fg'}{g^2}$ – **Quotient Rule**
5. $\dfrac{d}{dx}(c) = 0$
6. $\dfrac{d}{dx}(x^n) = n x^{n-1}$ – **Power Rule**
7. $\dfrac{d}{dx}(f(g(x))) = f'(g(x))\, g'(x)$

   This is the **Chain Rule**

Source : https://d20ohkaloyme4g.cloudfront.net/img/document_thumbnails/edea7c64fb478d3a5f53a7885af33724/thumb_1200_1553.png

Derivatives

Common Derivatives

$$\frac{d}{dx}(x)=1$$

$$\frac{d}{dx}(\sin x)=\cos x$$

$$\frac{d}{dx}(\cos x)=-\sin x$$

$$\frac{d}{dx}(\tan x)=\sec^2 x$$

$$\frac{d}{dx}(\sec x)=\sec x\tan x$$

$$\frac{d}{dx}(\csc x)=-\csc x\cot x$$

$$\frac{d}{dx}(\cot x)=-\csc^2 x$$

$$\frac{d}{dx}(\sin^{-1} x)=\frac{1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}(\cos^{-1} x)=-\frac{1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}(\tan^{-1} x)=\frac{1}{1+x^2}$$

$$\frac{d}{dx}(a^x)=a^x\ln(a)$$

$$\frac{d}{dx}(\mathbf{e}^x)=\mathbf{e}^x$$

$$\frac{d}{dx}(\ln(x))=\frac{1}{x},\quad x>0$$

$$\frac{d}{dx}(\ln|x|)=\frac{1}{x},\quad x\neq 0$$

$$\frac{d}{dx}(\log_a(x))=\frac{1}{x\ln a},\quad x>0$$

Source : https://d20ohkaloyme4g.cloudfront.net/img/document_thumbnails/edea7c64fb478d3a5f53a7885af33724/thumb_1200_1553.png

Derivatives

- Consider now the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, which is a real-valued function of n independent variables.
- We typically gather the variables into a vector x $(x_1 , x_2 , . . . , x_n)^T$
- the gradient of f at x: ($\partial f/\partial x_i$ represents the partial derivative of f with respect to x_i)

$$\nabla f(x) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix}$$

Derivatives

- The matrix of second partial derivatives of *f* is known as the Hessian

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Derivatives

- f is differentiable on a domain D if $\nabla f(x)$ exists for all $x \in D$
- f is twice differentiable on a domain D if $\nabla^2 f(x)$ exists for all $x \in D$
- f is continuously differentiable if $\nabla f(x)$ is a continuous functions of x
- f is twice continuously differentiable if $\nabla^2 f(x)$ is a continuous functions of x
- When f is twice continuously differentiable, the Hessian is a symmetric matrix

Example

$$f(x,y) = 4x^3y^2 - \sin(x)$$

$$\nabla f(x,y) = \begin{bmatrix} 12x^2y^2 - \cos(x) \\ 8x^3y \end{bmatrix}$$

$$\nabla f^2(x,y) = \begin{bmatrix} 24xy^2 + \sin(x) & 24x^2y \\ 24x^2y & 8x^3 \end{bmatrix}$$

Exercises

1. Given that v^T = [3, 5] and $f(x) = x^2 - 2|x|$, find f(v).
2. If D is a 3x3 matrix, proof that DD^T is symmetric.
3. If M is a 3x3 matrix and det(M) = n, find det(αM) where $\alpha \in \mathbb{R}$.
4. If v is a m elements column vector and A is a mXm matrix. Is $v^T A v = v A v^T$? Explain your answer.

5. Determine the range of value of *b* in the following matrix if it is a positive definite matrix.

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

6. Find the gradient and Hessian of the following function at (2,1):

$$f(x,y) = \ln(1+x+2y)$$