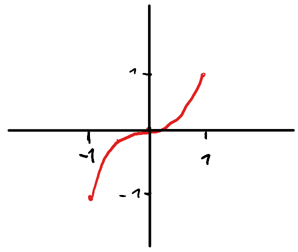
Functions and Domain Sketches

Sketch the following functions, curves, and domains:

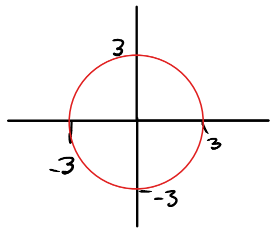
1. $f(x) = x^3, x \in [-1,1]$
2. $\{x | x < 1\}$
3. $\{(x,y) | x^2 + y^2 = 9\}$
4. $\{(x,y) | y = x - 2 \wedge x = 6 - y\}$

1.]

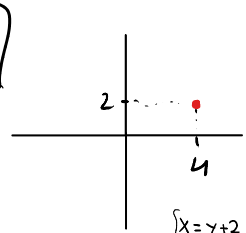


2.]

3.]



4.]



$\begin{cases} x = y+2 \\ x = 6-y \end{cases} \Rightarrow y+2 = 6-y \Downarrow y=2, x=4$

Multivariate Functions and their Derivatives

For each of the following functions $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$, compute their gradient ∇f_i and Hessian matrix H_f:

1. $f(x,y) = x^2 + y^2$
2. $f(x,y,z) = 2xy + x^2 + y^2 + z(x^2 - y^2)$
3. $f(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x}$, with $\mathbf{x} \in \mathbb{R}^n, n > 2$

1.] $\frac{\partial f}{\partial x} = 2x \quad \frac{\partial f}{\partial y} = 2y, \quad \nabla f(x,y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$

$$H_f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$x = (x_1, \dots, x_n)$

3] $\nabla f(x) = 2x = (2x_1, \dots, 2x_n)$

$$H_f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1 x_1} & \frac{\partial f}{\partial x_1 x_2} & \cdots \\ & \ddots & \\ \cdots & & \frac{\partial f}{\partial x_n x_n} \end{pmatrix} = 2 \cdot I_n$$

2.]

$$\nabla f(x,y,z) = \begin{bmatrix} 2y + 2x + 2xz \\ 2x + 2y - 2yz \\ x^2 - y^2 \end{bmatrix}$$

$$H_f(x,y,z) = \begin{pmatrix} 2+2z & 2 & 2x \\ 2 & 2-2z & -2y \\ 2x & -2y & 0 \end{pmatrix}$$

For each of the following functions $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^{m_i}$, compute their Jacobian matrix J_i:

1. $f(x,y,z) = (2xy + x^2 + y^2, \quad z(x^2 - y^2))$
2. $f(x,y) = \begin{pmatrix} x^2 y \\ x - y \\ 3xy^2 \end{pmatrix}$
3. $f(\mathbf{x}) = \mathbf{x} \odot \mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^n, n > 2$ and $\odot$ represents the element-wise multiplication.

1.] $$J_f(x,y,z) = \begin{pmatrix} 2y+2x & 2x+2y & 0 \\ 2zx & -2zy & x^2-y^2 \end{pmatrix}$$

2.] $J_{f(x,y)} = \begin{pmatrix} 2xy & x^2 \\ 1 & -1 \\ 3y^2 & 6xy \end{pmatrix}$ 3.] $f(x) = \begin{bmatrix} x_1 \cdot x_1 \\ \vdots \\ x_n \cdot x_n \end{bmatrix}$ $J_{f(x)} = \begin{pmatrix} 2x_1 \\ \vdots \\ 2x_n \end{pmatrix}$

Linear Systems

Solve the following linear systems:

1. $\begin{cases} x+y+z=2 \\ x-y-z=10 \\ x+2y-3z=5 \end{cases}$

2. $\begin{cases} x_1+2x_3-x_5=0 \\ 2x_2+2x_3-2x_5=1 \\ x_3+2x_4-x_5=2 \\ 2x_2+2x_3-x_5=-2 \\ x_4-2x_5=-1 \end{cases}$

1.] $\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 1 & -1 & -1 & 10 \\ 1 & 2 & -3 & 5 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & 0 & 0 & 12 \\ 0 & 1 & -4 & 3 \end{array}\right)$

$\begin{cases} x+y+z=2 \\ 2x=12 \\ y-4z=3 \end{cases} \Rightarrow \boxed{x=6}$

$\begin{cases} y+z=-4 \\ y-4z=3 \end{cases} \Rightarrow -4-z=3+4z$

$\Downarrow$

$-7=5z \Rightarrow \boxed{z=\frac{-7}{5}}$

$y=-4-z=-4+\frac{7}{5}=\frac{-\cancel{20}\,13}{5}$ $\boxed{y=\frac{-13}{5}}$

Consistent independent

2.

$\left(\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -1 & 0 \\ 0 & 2 & 2 & 0 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 & 2 \\ 0 & 2 & 2 & 0 & -1 & -2 \\ 0 & 0 & 0 & 1 & -2 & -1 \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -1 & 0 \\ 0 & 2 & 2 & 0 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 & -2 & -1 \end{array}\right)$

$\left.\begin{array}{r} x_1+2x_3-x_5=0 \\ 2x_2+2x_3-2x_5=1 \\ x_3+2x_4-x_5=2 \\ x_4-2x_5=-1 \\ x_5=-3 \end{array}\right\}$

$x_1=-2x_3+x_5 \Rightarrow \boxed{x_1=-29}$

$x_2=\frac{1-2x_3+2x_5}{2}=\frac{1-26-6}{2} \Rightarrow \boxed{x_2=\frac{-31}{2}}$

$x_3=2-2x_4+x_5=2+14-3 \Rightarrow \boxed{x_3=13}$

$x_4=-1+2x_5 \Rightarrow \boxed{x_4=-7}$

$\boxed{x_5=-3}$

Consistent independent

Engineering example

Suppose we use the following model with unknown parameters to describe a specific chemical process.

$$y(x)=\beta_0 e^x+\beta_1 sin(\tfrac{2\pi}{3}x)+\beta_2 x^3,$$

where β_i - are the parameters. To perform data fitting and roughly estimate the parameters of the model, we use linear regression over a small set of inputs $\{x_i\}_{i=1}^4$, and corresponding outputs $\{y_i\}_{i=1}^4$.

i	x_i	y_i
1	−1	−5
2	0	1
3	1	6
4	2	34

1. Rewrite the model equation in a format $y(x)=\mathbf{a}(x)*\mathbf{X}$, where $\mathbf{X}=\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}$ is a column vector of unknown parameters, and $\mathbf{a}(x)$ - is a row vector of corresponding functions.

$a(x)=\left[e^x \quad \sin\left(\frac{2\pi}{3}x\right) \quad x^3\right]$

$y(x)=a(x)*X$

2. Using the rewritten model equation and the provided data, formulate an overdetermined linear system $\mathbf{Y} = \mathbf{A} * \mathbf{X}$, where $\mathbf{Y} = \begin{bmatrix} .. \\ y_i \\ .. \end{bmatrix}$ is a vector of outputs of the model, and $\mathbf{A} = \begin{bmatrix} \cdots\cdots\cdots\cdots \\ \text{---} \quad \mathbf{a}(x_i) \quad \text{---} \\ \cdots\cdots\cdots\cdots \end{bmatrix}$ is a model matrix, formed of the above-mentioned row vectors, evaluated for each given input. Note that since the matrix represents the model evaluated for the fixed dataset, the x dependence is omitted. $\mathbf{X}$ is the above-mentioned vector of parameters, and $*$ represents a standard matrix-vector product.

$$\overbrace{\begin{pmatrix} e^{-1} & -\frac{\sqrt{3}}{2} & -1 \\ 1 & 0 & 0 \\ e & \frac{\sqrt{3}}{2} & 1 \\ e^2 & -\frac{\sqrt{3}}{2} & 8 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} B_0 \\ B_1 \\ B_2 \end{pmatrix}}^{X} = \overbrace{\begin{pmatrix} -5 \\ 1 \\ 6 \\ 34 \end{pmatrix}}^{Y}$$

3. Since the problem is overdetermined; a proposed solution technique is to optimally satisfy the data in the 2-norm sense. Algebraically put, the problem is as follows.

$$min||\mathbf{Y} - \mathbf{A} * \mathbf{X}||_2^2$$

Expanding the definition of the 2-norm, show that this formulation minimizes the quadratic objective function.

$$||Y - AX||_2^2 = (Y-Ax)^T(Y-Ax) = (Y^T - X^TA^T)(Y - Ax) = Y^TY - Y^TAX - X^TA^TY + X^TA^TAX =$$

$$= X^TA^TAX - Y^TAX - \underbrace{X^TA^TY}_{\substack{\text{is a scalar} \\ \text{so we can} \\ \text{transpose } X^TA^TY = Y^TAX}} + Y^TY = \underbrace{X^TA^TAX}_{} - 2Y^TAX + Y^TY$$

↳ It is a quadratic objective function.

4. Compute the gradient of the resulting function.

$$\nabla||Y-AX||_2^2 = (\underbrace{A^TA}_{\text{symetric} \Rightarrow A^TA = (A^TA)^T} + (A^TA)^T) \cdot X - 2Y^TA \Rightarrow 2A^TAX - 2Y^TA$$

5. Compute the Hessian of the corresponding function.

$$\nabla^2||Y-AX||_2^2 = \underbrace{2A^TA}_{\text{Symmetric}}$$

6. Is the Hessian symmetric positive-definite? Check by computing the eigenvalue decomposition of the Hessian. (*Hint.* For symmetric matrices, positive (negative) definiteness means that all of the eigenvalues are strictly positive (negative). If, however, at least one of the eigenvalues is zero, but the rest have the same sign, the matrix is called positive (negative) semi-definite. If none of the above holds, the matrix is called indefinite.)

We can say that the Hessian is positive definite because it has the form $2A^TA$ and therefore all the eigenvalues are positive.

Proof: Let λ be an eigenvalue and v an eigenvector:

$A^T A v = \lambda v$. If we multiply on the left by v^T

$v^TA^TAv = \lambda v^Tv$

$(Av)^T(Av) = \lambda v^Tv \Rightarrow \overbrace{||Av||_2^2}^{\geq 0} = \lambda \overbrace{||v||_2^2}^{\geq 0} \Rightarrow \lambda \geq 0$