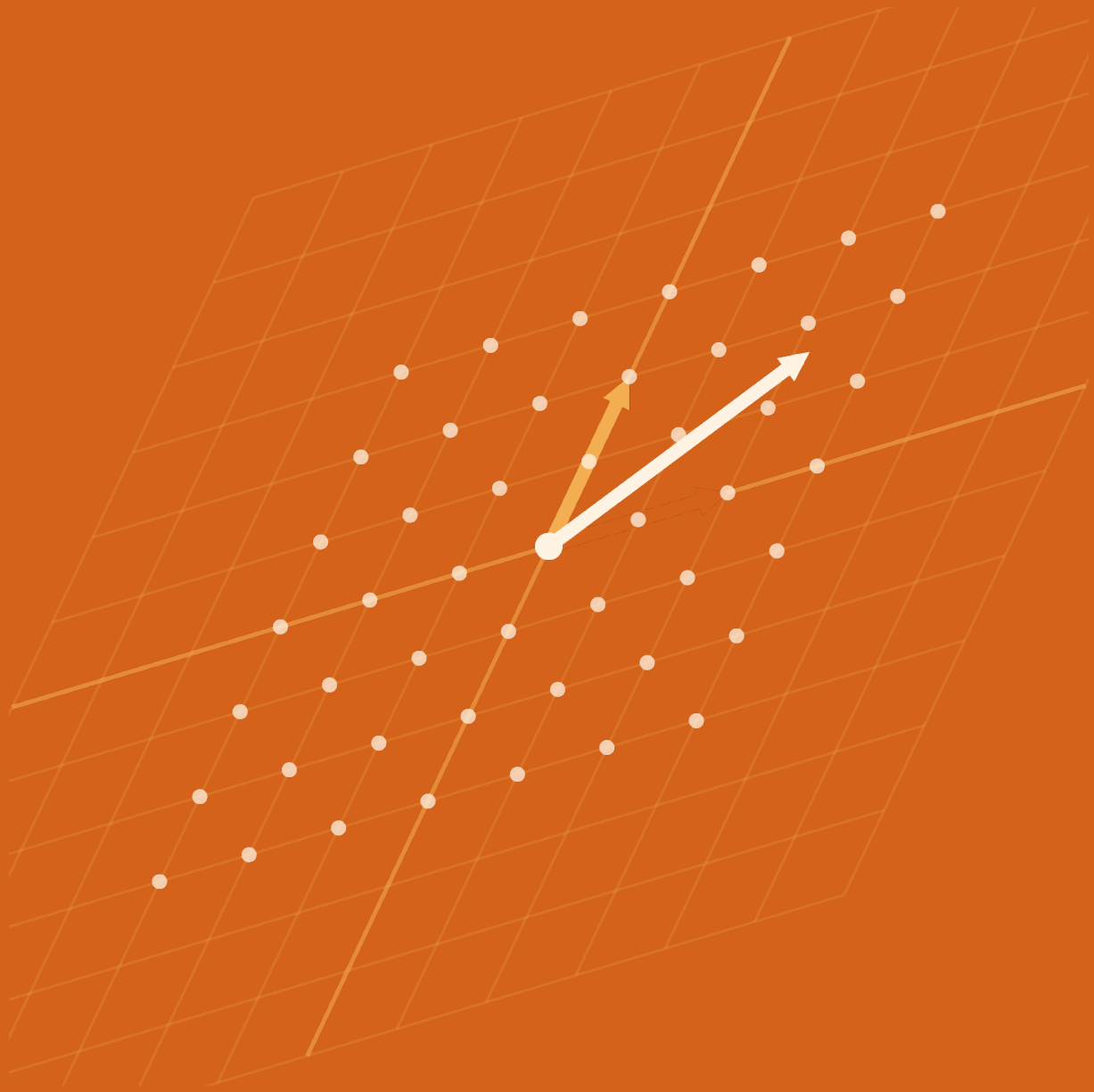
LINEAR ALGEBRA FOR AIML AND 6G

Vector and Matrix Calculus



Lecture 20

Prof. Aditya K. Jagannatham
IIT Kanpur

Outline

1. **Vector derivatives**
   Derivatives of scalars and vectors with respect to vectors
2. **Jacobian**
   Jacobian of a transformation, with spherical coordinates
3. **Chain rule**
   Derivatives of composed vector functions
4. **Matrix derivatives**
   Derivatives of the trace and the determinant

01

Vector derivatives

Derivatives of scalars and vectors with respect to vectors



Vector calculus

- Consider vector $\bar{\mathbf{x}}$

$$\bar{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Vector calculus

- Derivative of a scalar function $y = y(\bar{\mathbf{x}})$ with respect to vector $\bar{\mathbf{x}}$ is

$$\frac{\partial y}{\partial \bar{\mathbf{x}}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}$$

- This is also denoted by ∇y

Vector calculus

- Consider x to be a scalar
- Derivative of a **vector function** $\bar{\mathbf{y}}$ with respect to x is

$$\frac{\partial \bar{\mathbf{y}}}{\partial x} = \left[\frac{\partial y_1}{\partial x} \ \frac{\partial y_2}{\partial x} \ \cdots \ \frac{\partial y_m}{\partial x}\right]$$

Vector calculus

- Derivative of a vector $\bar{\mathbf{y}}$ with respect to another vector $\bar{\mathbf{x}}$ is

$$\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Vector calculus – example

- Consider

$$\bar{\mathbf{y}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1^3 + 2x_2 \\ 4x_3 - x_2^2 \end{bmatrix}$$

$$\bar{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Vector calculus – example

- Derivative of $\bar{\mathbf{y}}$ with respect to $\bar{\mathbf{x}}$ is

$$\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} \\ \frac{\partial y_1}{\partial x_3} & \frac{\partial y_2}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 3x_1^2 & 0 \\ 2 & -2x_2 \\ 0 & 4 \end{bmatrix}$$

Calculus results

- Consider

$$y = \bar{\mathbf{c}}^T \bar{\mathbf{x}} = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n$$
$$= \bar{\mathbf{x}}^T \bar{\mathbf{c}}$$

Calculus results

- Then,

$$\frac{\partial y}{\partial \bar{\mathbf{x}}} = \frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{c}}^T \bar{\mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \bar{\mathbf{c}}$$

Calculus results

- Consider now

$$y = \bar{\mathbf{x}}^T \mathbf{A} \bar{\mathbf{x}} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} x_j$$

- The derivative with respect to $\bar{\mathbf{x}}$ is obtained as

$$\frac{\partial y}{\partial \bar{\mathbf{x}}} = \frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{x}}^T (\mathbf{A} \bar{\mathbf{x}}) + \frac{\partial}{\partial \bar{\mathbf{x}}} \left(\bar{\mathbf{x}}^T \mathbf{A} \right) \bar{\mathbf{x}}$$

$$= \mathbf{A} \bar{\mathbf{x}} + \left(\bar{\mathbf{x}}^T \mathbf{A} \right)^T = \mathbf{A} \bar{\mathbf{x}} + \mathbf{A}^T \bar{\mathbf{x}}$$

Calculus results

- When $\mathbf{A} = \mathbf{A}^T$

$$\frac{\partial y}{\partial \bar{\mathbf{x}}} = \mathbf{A}\bar{\mathbf{x}} + \mathbf{A}^T\bar{\mathbf{x}} = 2\mathbf{A}\bar{\mathbf{x}}$$

02

Jacobian

Jacobian of a transformation, with spherical coordinates



Jacobian

- In **multivariate analysis**, if $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are of the same order
- Consider the transformation $\bar{\mathbf{y}} = \bar{\mathbf{y}}(\bar{\mathbf{x}})$

Jacobian

- The **Jacobian** of this transformation is

$$J = \left| \left(\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} \right)^T \right| = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_n} \end{vmatrix}$$

Jacobian – example

- Transformation from spherical to **Cartesian coordinates** is defined by

$$x = r\sin\theta\cos\phi = x\left(r,\theta,\phi\right)$$

$$y = r\sin\theta\sin\phi = y\left(r,\theta,\phi\right)$$

$$z = r\cos\theta = z\left(r,\theta,\phi\right)$$

$$r > 0, 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$$

Jacobian – example

- **Jacobian** of the transformation from spherical to Cartesian coordinates

$$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix}$$

Jacobian – example

- Jacobian of the transformation from spherical to Cartesian coordinates is

$$\begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}$$

Jacobian – example

- The Jacobian can be simplified as

$$= \cos\theta \left(r^2 \cos\theta \sin\theta \cos^2\phi + r^2 \cos\theta \sin\theta \sin^2\phi\right)$$

$$+ r \sin\theta \left(r \sin^2\theta \cos^2\phi + r \sin^2\theta \sin^2\phi\right)$$

$$= r^2 \cos^2\theta \sin\theta + r^2 \sin^3\theta = r^2 \sin\theta$$

03

Chain rule

Derivatives of composed vector functions



Chain rule

- Let

$$\bar{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \bar{\mathbf{y}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \bar{\mathbf{z}} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_r \end{bmatrix}$$

Chain rule

- Then

$$\left[\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{x}}}\right]_{ji} = \frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k}\frac{\partial y_k}{\partial x_j}$$

$$= \sum_{k=1}^{m} \left[\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{y}}}\right]_{ki} \left[\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}}\right]_{jk}$$

Chain rule

- This can be written as

$$\left[\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{x}}}\right]_{ji} = \left[\left(\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{x}}}\right)^T\right]_{ij}$$

$$= \sum_{k=1}^{m} \left[\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{y}}}\right]_{ki} \left[\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}}\right]_{jk} = \sum_{k=1}^{m} \left[\left(\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{y}}}\right)^T\right]_{ik} \left[\left(\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}}\right)^T\right]_{kj}$$

Chain rule

- Therefore, the **chain rule** follows as

$$\left(\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{x}}}\right)^T = \left(\frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{y}}}\right)^T \left(\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}}\right)^T$$

$$= \left(\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} \frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{y}}}\right)^T$$

$$\Rightarrow \frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{x}}} = \frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} \frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{y}}}$$

Chain rule

- Let $\bar{\mathbf{w}}$ be a function of $\bar{\mathbf{z}}$
- One can extend the chain rule as

$$\frac{\partial \bar{\mathbf{w}}}{\partial \bar{\mathbf{x}}} = \frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} \frac{\partial \bar{\mathbf{z}}}{\partial \bar{\mathbf{y}}} \frac{\partial \bar{\mathbf{w}}}{\partial \bar{\mathbf{z}}}$$

04

Matrix derivatives

Derivatives of the trace and the determinant



Matrix derivatives

- Let $\mathbf{X}$ be a matrix of size $m \times n$
- Let $y = f(\mathbf{X})$ be a scalar function of a matrix $\mathbf{X}$.

Matrix derivatives

- The derivative of y with respect to $\mathbf{X}$ is

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1n}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{m1}} & \frac{\partial y}{\partial x_{m2}} & \cdots & \frac{\partial y}{\partial x_{mn}} \end{bmatrix}$$

Derivative of trace

- Consider the trace function defined as

$$y = \mathrm{tr}\left(\mathbf{X}\right) = \sum_{i=1}^{n} x_{ii}$$

$$= x_{11} + x_{22} + \ldots + x_{nn}$$

Derivative of trace

- Observe

$$\frac{\partial}{\partial x_{ii}} \mathrm{tr}\left(\mathbf{X}\right) = 1$$

$$\frac{\partial}{\partial x_{ij}} \mathrm{tr}\left(\mathbf{X}\right) = 0, i \neq j$$

$$\Rightarrow \frac{\partial}{\partial \mathbf{X}} \mathrm{tr}\left(\mathbf{X}\right) = \mathbf{I}$$

Derivative of determinant

- Consider

$$y = |\mathbf{X}|$$

- $(\partial/\partial\mathbf{X})|\mathbf{X}|$ can be determined as follows

Derivative of determinant

- Note that

$$|\mathbf{X}| = x_{i1}C_{i1} + x_{i2}C_{i2} + \ldots + x_{in}C_{in} = \sum_{j=1}^{n} x_{ij}C_{ij}$$

$$\Rightarrow \frac{\partial}{\partial x_{ij}} |\mathbf{X}| = C_{ij}$$

Derivative of determinant

- Therefore

$$\frac{\partial}{\partial \mathbf{X}} |\mathbf{X}| = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{bmatrix} = \mathbf{C}$$

- where C_{ij} is the **cofactor** of x_{ij}

Derivative of determinant

- Recall

$$\mathbf{X}^{-1} = \frac{\mathbf{C}^T}{|\mathbf{X}|} \Rightarrow \mathbf{C} = \left(|\mathbf{X}|\,\mathbf{X}^{-1}\right)^T$$

- Therefore,

$$\frac{\partial}{\partial \mathbf{X}} |\mathbf{X}| = \mathbf{C} = \left(|\mathbf{X}|\,\mathbf{X}^{-1}\right)^T$$

Summary

- The gradient $\partial y/\partial \bar{\mathbf{x}}$ collects the partial derivatives; $\partial \bar{\mathbf{y}}/\partial \bar{\mathbf{x}}$ has entries $(\partial y_j)/(\partial x_i)$
- $(\partial/\partial \bar{\mathbf{x}})\bar{\mathbf{c}}^T\bar{\mathbf{x}} = \bar{\mathbf{c}}$ and $(\partial/\partial \bar{\mathbf{x}})\bar{\mathbf{x}}^T\mathbf{A}\bar{\mathbf{x}} = (\mathbf{A} + \mathbf{A}^T)\bar{\mathbf{x}}$, which is $2\mathbf{A}\bar{\mathbf{x}}$ for symmetric $\mathbf{A}$
- The **Jacobian** is the determinant of the derivative matrix; for spherical coordinates it is $r^2 \sin\theta$
- **Chain rule**: $(\partial \bar{\mathbf{z}}/\partial \bar{\mathbf{x}}) = (\partial \bar{\mathbf{y}}/\partial \bar{\mathbf{x}})\partial \bar{\mathbf{z}}/\partial \bar{\mathbf{y}}$
- $(\partial/\partial \mathbf{X})\mathrm{tr}(\mathbf{X}) = \mathbf{I}$ and $(\partial/\partial \mathbf{X})|\mathbf{X}| = (|\mathbf{X}|\mathbf{X}^{-1})^T$



Thank You!

Questions?