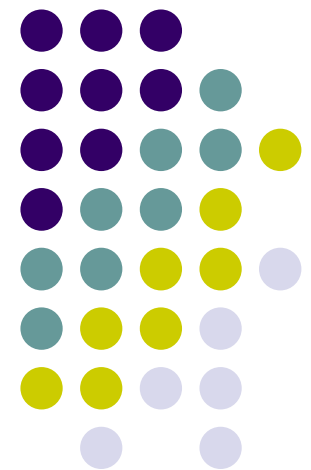


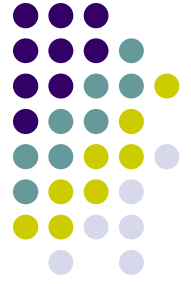
ME751

Advanced Computational Multibody Dynamics

Review: Elements of Linear Algebra & Calculus
September 9, 2016

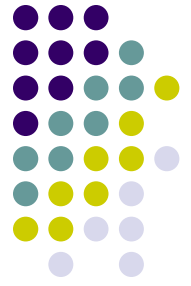


Quote of the day



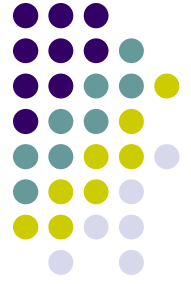
If you can't convince them, confuse them.
- Harry S. Truman (US President)

Before we get started...

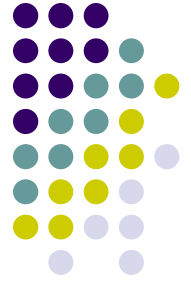


- Last Time:
 - Class Intro + Syllabus Outline
- Today:
 - Review of elements of Linear Algebra
 - Review of elements of Calculus (two definitions and three theorems)
- Purpose of today's class
 - Not introducing any new concepts, but rather *zipping through* a collection of concepts that you learned in the past and are going to be used time and again in ME751
 - An enumeration of things good to know/understand
 - I expect that you'll go through these slides and make sure it all makes sense to you

Notation Conventions



- A bold upper case letter denotes matrices
 - Example: **A**, **B**, etc.
- A bold lower case letter denotes a vector
 - Example: **v**, **s**, etc.
- A letter in italics format denotes a scalar quantity
 - Example: *a*, *b*



Matrix Review

- Matrix: a tableau of elements

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1^T \\ \boldsymbol{\alpha}_2^T \\ \dots \\ \boldsymbol{\alpha}_m^T \end{bmatrix}$$

- Matrix addition:

$$\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n$$

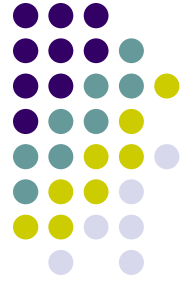
$$\mathbf{B} = [b_{ij}] \in \mathbb{R}^{m \times n}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n$$

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = [c_{ij}] \in \mathbb{R}^{m \times n}, \quad c_{ij} = a_{ij} + b_{ij}$$

- Addition is commutative

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

Matrix Multiplication



- Dimension constraints on matrices so that they can be multiplied:
 - # of columns of first matrix is equal to # of rows of second matrix

$$\mathbf{A} = [a_{ij}], \quad \mathbf{A} \in \mathbb{R}^{m \times n}$$

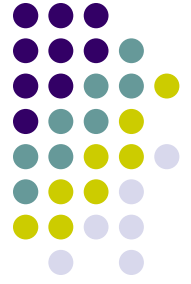
$$\mathbf{C} = [c_{ij}], \quad \mathbf{C} \in \mathbb{R}^{n \times p}$$

$$\mathbf{D} = \mathbf{A} \cdot \mathbf{C} = [d_{ij}], \quad \mathbf{D} \in \mathbb{R}^{m \times p}$$

$$d_{ij} = \sum_{k=1}^n a_{ik} c_{kj}$$

- This operation is not commutative
- Distributivity of matrix multiplication with respect to matrix addition:

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{C}$$



Matrix-Vector Multiplication

- A **column-wise** perspective on matrix-vector multiplication

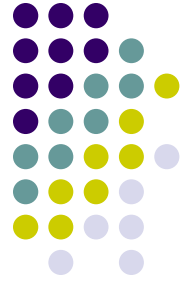
$$\mathbf{A}\mathbf{v} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix} = \sum_{i=1}^n v_i \mathbf{a}_i$$

- Example:

$$\mathbf{A}\mathbf{v} = \begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 3 & 1 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 1 & -1 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} \cdot (1) + \begin{bmatrix} 4 \\ 3 \\ 0 \\ 1 \end{bmatrix} \cdot (2) + \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix} \cdot (-1) + \begin{bmatrix} 0 \\ 1 \\ -1 \\ -2 \end{bmatrix} \cdot (1) = \begin{bmatrix} 7 \\ 8 \\ -3 \\ 1 \end{bmatrix}$$

- A **row-wise** perspective on matrix-vector multiplication: $\mathbf{A}\mathbf{v} = \begin{bmatrix} \boldsymbol{\alpha}_1^T \\ \boldsymbol{\alpha}_2^T \\ \dots \\ \boldsymbol{\alpha}_m^T \end{bmatrix} \mathbf{v} = \begin{bmatrix} \boldsymbol{\alpha}_1^T \mathbf{v} \\ \boldsymbol{\alpha}_2^T \mathbf{v} \\ \dots \\ \boldsymbol{\alpha}_m^T \mathbf{v} \end{bmatrix}$

Matrix Review [Cntd.]



- Scaling of a matrix by a real number: scale each entry of the matrix

$$\alpha \cdot \mathbf{A} = \alpha \cdot [a_{ij}] = [\alpha \cdot a_{ij}]$$

- Example:

$$(1.5) \cdot \begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 3 & 1 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 1 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1.5 & 6 & 3 & 0 \\ 3 & 4.5 & 1.5 & 1.5 \\ -1.5 & 0 & 1.5 & -1.5 \\ 0 & 1.5 & -1.5 & -3 \end{bmatrix}$$

- Transpose of a matrix \mathbf{A} dimension $m \times n$: a matrix $\mathbf{B} = \mathbf{A}^T$ of dimension $n \times m$ whose (i, j) entry is the (j, i) entry of original matrix \mathbf{A}

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 3 & 1 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 1 & -1 & -2 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 4 & 3 & 0 & 1 \\ 2 & 1 & 1 & -1 \\ 0 & 1 & -1 & -2 \end{bmatrix}$$



Linear Independence of Vectors

- Definition: linear independence of a set of m vectors, $\mathbf{v}_1, \dots, \mathbf{v}_n$:

$$\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^m$$

- The vectors are linearly independent if the following condition holds

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n = \mathbf{0} \quad \Rightarrow \quad \alpha_1 = \dots = \alpha_n = 0$$

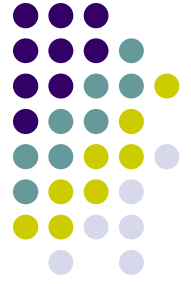
- If a set of vectors are not linearly independent, they are called dependent

- Example:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -3 \\ -6 \end{bmatrix}$$

- Note that $\mathbf{v}_1 - 2\mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$

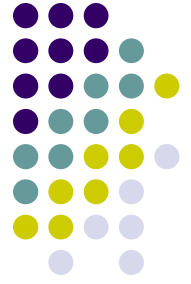
Matrix Rank



- Row rank of a matrix
 - Largest number of rows of the matrix that are linearly independent
 - A matrix is said to have full row rank if the rank of the matrix is equal to the number of rows of that matrix
- Column rank of a matrix
 - Largest number of columns of the matrix that are linearly independent
- Important results
 - For any matrix, the row rank and column rank are the same
 - This number is simply called the rank of the matrix
 - It follows that

$$\text{rank}(C) = \text{rank}(C^T)$$

Matrix Rank, Example

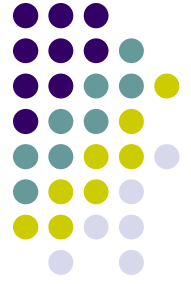


- What is the row rank of the matrix \mathbf{J} ?

$$\mathbf{J} = \begin{bmatrix} 2 & 1 & -1 & 0 \\ 4 & -2 & -2 & 1 \\ 0 & -4 & 0 & 1 \end{bmatrix}$$

- What is the rank of \mathbf{J} ?

Matrix & Vector Norms



- Norm of a vector
 - Definitions: norm 1, norm 2 (or Euclidian), and Infinity norm

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \quad \|\mathbf{x}\|_\infty = \max |x_i|$$

- Norm of a matrix (the “consistent form” – there are several other norms)
 - Definition: norm 1, norm 2 (or Euclidian), and Infinity

$$\|\mathbf{A}\|_p = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p}$$

$$\|\mathbf{A}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

$$\|\mathbf{A}\|_2 = \sqrt{\rho(\mathbf{A}^T \mathbf{A})}$$

$$\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

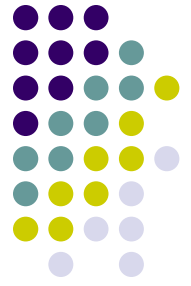
Matrix & Vector Norms, Example



- Find norm 1, Euclidian, and Infinity for the following matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix}$$

Matrix Review [Cntd.]



- Symmetric matrix: a square matrix \mathbf{A} for which $\mathbf{A}=\mathbf{A}^T$
- Skew-symmetric matrix: a square matrix \mathbf{B} for which $\mathbf{B}=-\mathbf{B}^T$
- Examples:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 3 \\ -1 & 3 & 4 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 4 \\ -2 & -4 & 0 \end{bmatrix}$$

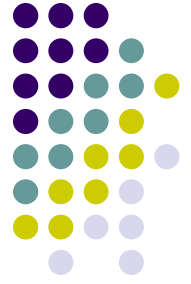
- Singular matrix: square matrix whose determinant is zero

$$\det(\mathbf{A}) = 0, \qquad \mathbf{A} \in \mathbb{R}^{n \times n}$$

- Inverse of a square matrix \mathbf{A} : a matrix of the same dimension, called \mathbf{A}^{-1} , that satisfies the following:

$$\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n, \qquad \mathbf{A} \in \mathbb{R}^{n \times n}$$

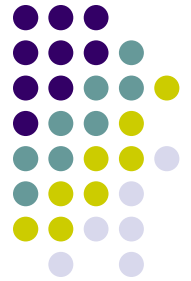
Singular vs. Nonsingular Matrices



- Let \mathbf{A} be a square matrix of dimension n . The following are equivalent:
 - $\mathbf{Ax} = \mathbf{b}$ has a unique solution for any $\mathbf{b} \in \mathbb{R}^n$.
 - $\mathbf{Ax} = \mathbf{b}$ has a solution for any $\mathbf{b} \in \mathbb{R}^n$.
 - $\mathbf{Ax} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$.
 - \mathbf{A}^{-1} exists.
 - $\text{Determinant}(\mathbf{A}) \neq 0$.
 - $\text{rank}(\mathbf{A}) = n$.

Orthogonal & Orthonormal Matrices

[we'll work w/ a lot of orthonormal matrices]



- Definition (\mathbf{Q} , orthogonal matrix): a square matrix \mathbf{Q} is orthogonal if the product $\mathbf{Q}^T\mathbf{Q}$ is a diagonal matrix
- Matrix \mathbf{Q} is called orthonormal if it's orthogonal and also $\mathbf{Q}^T\mathbf{Q}=\mathbf{I}_n$
 - Note that people in general don't make a distinction between an orthogonal and orthonormal matrix
- Note that if \mathbf{Q} is an orthonormal matrix, then $\mathbf{Q}^{-1}=\mathbf{Q}^T$
- Example, orthonormal matrix:

$$\mathbf{Q} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{bmatrix}$$



Remark:

On the Columns of an Orthonormal Matrix

- Assume \mathbf{Q} is an orthonormal matrix

$$\mathbf{Q} \in \mathbb{R}^{n \times n} \quad \mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_n] \quad \leftarrow \text{orthonormal}$$

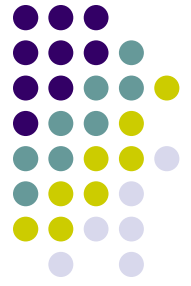
$$\mathbf{Q}^T \mathbf{Q} = I \quad \Rightarrow \quad \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} [\mathbf{q}_1, \dots, \mathbf{q}_n] = \begin{bmatrix} \mathbf{q}_1^T \mathbf{q}_1 & \dots & \mathbf{q}_1^T \mathbf{q}_n \\ \vdots & \ddots & \vdots \\ \mathbf{q}_n^T \mathbf{q}_1 & \dots & \mathbf{q}_n^T \mathbf{q}_n \end{bmatrix}$$



$$\mathbf{q}_i^T \cdot \mathbf{q}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

- In other words, the columns (and the rows) of an orthonormal matrix have unit norm and are mutually perpendicular to each other

Condition Number of a Matrix

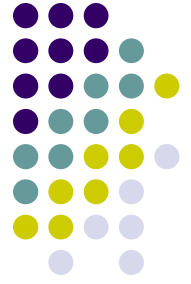


- Let \mathbf{A} be a square matrix. By definition, its condition number is

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$$

- Note that condition number depends on the norm used in its evaluation
- The concept of ill-conditioned linear system $\mathbf{Ax}=\mathbf{b}$:
 - A system for which small perturbations in \mathbf{b} lead to large changes in solution \mathbf{x}
 - NOTE: A linear system is ill-condition if $\text{cond}(\mathbf{A})$ is large
- Three quick remarks:
 - The closer a matrix is to being singular, the larger its condition number
 - You can't get $\text{cond}(\mathbf{A})$ to be smaller than 1
 - If \mathbf{Q} is orthonormal, then $\text{cond}_2(\mathbf{Q})=1$

Let's flex our brain muscles

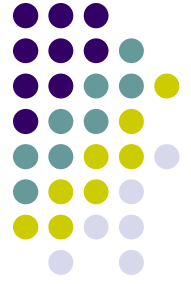


- Show that

$$\text{cond}_2(\mathbf{Q}) = 1$$

Condition Number of a Matrix

Example



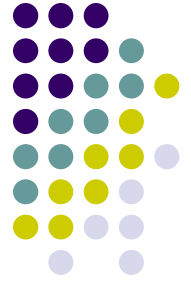
$$\begin{cases} 7x_1 + 10x_2 = b_1 \\ 5x_1 + 7x_2 = b_2 \end{cases}$$

$$\mathbf{A} = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} \quad \mathbf{A}^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}$$

$$\text{cond}(\mathbf{A})_1 = \|\mathbf{A}\|_1 \cdot \|\mathbf{A}^{-1}\|_1 = 289$$

$$\text{cond}(\mathbf{A})_2 = \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 \approx 223$$

$$\text{cond}(\mathbf{A})_\infty = \|\mathbf{A}\|_\infty \cdot \|\mathbf{A}^{-1}\|_\infty = 289$$



Other Useful Formulas

- If **A** and **B** are invertible, their product is invertible and

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

- Also,

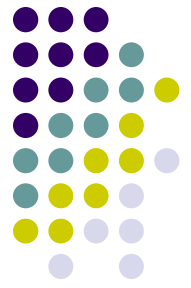
$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$

- For any two matrices **A** and **B** that can be multiplied

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$$

- For any three matrices **A**, **B**, and **C** that can be multiplied

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$



Lagrange Multiplier Theorem

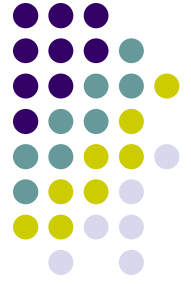
- Theorem:

Assume that a vector $\mathbf{b} \in \mathbb{R}^n$, and a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, with $m < n$, are two quantities related by the following relationship: **ANY** vector $\mathbf{x} \in \mathbb{R}^n$ that is perpendicular on the rows on \mathbf{A} is also perpendicular on \mathbf{b} ; i.e., $\mathbf{x}^T \mathbf{b} = 0$ as soon as $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Then it turns out that in fact \mathbf{b} is a linear combination of the rows of \mathbf{A} . In other words, there is a so called “Lagrange Multiplier” λ such that $\mathbf{b} = -\mathbf{A}^T \lambda$, or equivalently, $\mathbf{b} + \mathbf{A}^T \lambda = \mathbf{0}$.

[(Ex. 6.3.3) – Haug's Book]

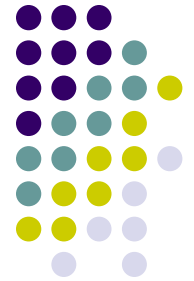
Example: Lagrange Multipliers



$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 3 & 1 \end{bmatrix}$$

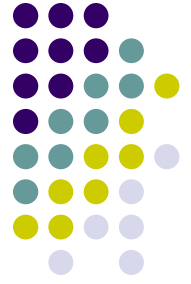
$$\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

- First, show that any for any $\mathbf{x}=[x_1 \ x_2 \ x_3]^T$, one has that $\mathbf{x}^T \mathbf{b}=0$ as soon as $\mathbf{Ax}=\mathbf{0}$
- Next, show that there is indeed a vector λ such that $\mathbf{b} + \mathbf{A}^T \lambda = \mathbf{0}$



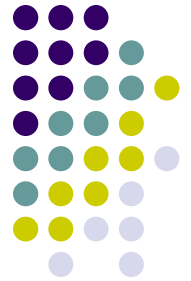
End: Review of Linear Algebra
Begin: Review of Calculus

Derivatives of Functions



- GOAL: Understand how to
 - Take time derivatives of vectors and matrices
 - Take partial derivatives of a function with respect to its arguments
 - We will use a matrix-vector notation for computing these partial derivs.
 - Taking partial derivatives might be challenging in the beginning
 - The use of partial derivatives is a recurring theme in the literature
- Time and partial derivatives: this horse has been beaten to death in ME451

Taking time derivatives of a time dependent vector



- FRAMEWORK:

- Vector \mathbf{r} is represented as a function of time, and it has three components: $x(t)$, $y(t)$, $z(t)$:

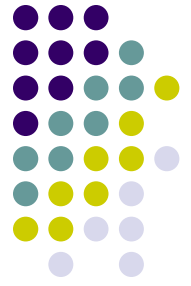
$$\mathbf{r}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}$$

- Its components change, but the vector is represented in a **fixed** reference frame

- THEN:

$$\dot{\mathbf{r}}(t) = \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{bmatrix}, \quad \ddot{\mathbf{r}}(t) = \begin{bmatrix} \ddot{x}(t) \\ \ddot{y}(t) \\ \ddot{z}(t) \end{bmatrix}, \quad \text{etc.}$$

Time Derivatives, Vector Related Operations



- Assume that $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^n$ depend on time. Then it can be proved that the following hold:

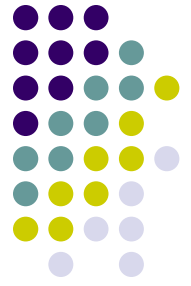
$$\frac{d}{dt}(\alpha \mathbf{a}) = \frac{d\alpha}{dt} \mathbf{a} + \alpha \frac{d\mathbf{a}}{dt} = \dot{\alpha} \mathbf{a} + \alpha \dot{\mathbf{a}}$$

$$\frac{d}{dt}(\mathbf{a} + \mathbf{b}) = \frac{d\mathbf{a}}{dt} + \frac{d\mathbf{b}}{dt} = \dot{\mathbf{a}} + \dot{\mathbf{b}}$$

$$\frac{d}{dt}(\mathbf{a}^T \mathbf{b}) = \frac{d\mathbf{a}^T}{dt} \mathbf{b} + \mathbf{a}^T \frac{d\mathbf{b}}{dt} = \dot{\mathbf{a}}^T \mathbf{b} + \mathbf{a}^T \dot{\mathbf{b}}$$

$$\mathbf{a}^T \mathbf{a} = \text{const} \quad \Rightarrow \quad \mathbf{a}^T \dot{\mathbf{a}} = 0$$

Taking time derivatives of MATRICES



- By definition, the time derivative of a matrix is obtained by taking the time derivative of each entry in the matrix
- Simple extension of what seen for vector derivatives
- Assume that $\alpha \in \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times n}$, and $\mathbf{C} \in \mathbb{R}^{n \times p}$ depend on time. Then it can be proved that the following hold:

$$\frac{d}{dt}(\alpha \mathbf{A}) = \frac{d\alpha}{dt} \mathbf{A} + \alpha \frac{d\mathbf{A}}{dt} = \dot{\alpha} \mathbf{A} + \alpha \dot{\mathbf{A}}$$

$$\frac{d}{dt}(\mathbf{A} + \mathbf{B}) = \frac{d\mathbf{A}}{dt} + \frac{d\mathbf{B}}{dt} = \dot{\mathbf{A}} + \dot{\mathbf{B}}$$

$$\frac{d}{dt}(\mathbf{A}\mathbf{C}) = \frac{d\mathbf{A}}{dt} \mathbf{C} + \mathbf{A} \frac{d\mathbf{C}}{dt} = \dot{\mathbf{A}}\mathbf{C} + \mathbf{A}\dot{\mathbf{C}}$$

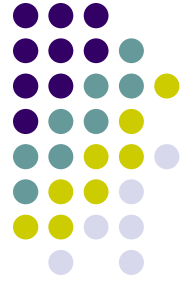


Done with Time Derivatives

...

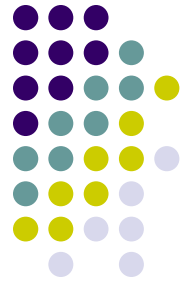
Moving on to Partial Derivatives

Derivatives of Functions: Why Bother?



- Partial derivatives are essential in this class
 - In computing the Jacobian matrix associated with the constraints that define the joints present in a mechanism
 - Essential in computing the Jacobian matrix of any nonlinear system that you will have to solve when using implicit integration to find the time evolution of a dynamic system
- Beyond this class
 - Whenever you do a sensitivity analysis (in optimization, for instance) you need partial derivatives of your functions

What's the story behind the concept of partial derivative?



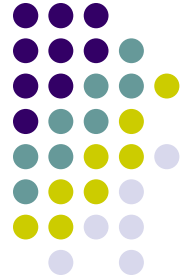
- What's the meaning of a partial derivative?
 - It captures the “sensitivity” of a function with respect to a variable the function depends upon
 - Shows how much the function changes when the variable changes a bit
- Simplest case of partial derivative: you have one function that depends on one variable:

$$f(x) = \ln x \quad , \quad g(z) = \sin(4z + \pi) \quad , \quad \text{etc.}$$

- Then,

$$\frac{\partial f}{\partial x} = \frac{1}{x} \quad , \quad \frac{\partial g}{\partial z} = 4 \cos(4z + \pi) \quad , \quad \text{etc.}$$

Partial Derivative, Two Variables



- Suppose you have one function but it depends on **two** variables, say x and y :

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

- To simplify the notation, an array \mathbf{q} is introduced:

$$\mathbf{q} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$$

- With this, the partial derivative of $f(\mathbf{q})$ wrt \mathbf{q} is defined as

Notation...

$$\frac{\partial f}{\partial \mathbf{q}} = \boxed{f_{\mathbf{q}}} = \left[\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right] = [2x \cos(x^2 + 3y^2) \quad 6y \cos(x^2 + 3y^2)]$$

...and here is as good as it gets (vector function)



- You have a group of “ m ” functions that are gathered together in an array, and they depend on a collection of “ n ” variables:

f_1, f_2, \dots, f_m depend on x_1, x_2, \dots, x_n

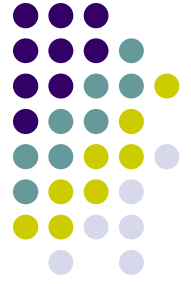
- The array that collects all “ m ” functions is called \mathbf{F} :

$$\mathbf{F}(x_1, x_2, \dots, x_n) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \dots \\ f_m(x_1, x_2, \dots, x_n) \end{bmatrix} \in \mathbb{R}^m$$

- The array that collects all “ n ” variables is called \mathbf{q} :

$$\mathbf{q} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

Most general partial derivative (Vector Function, cntd)



- Then, in the most general case, by definition

$$\frac{\partial \mathbf{F}}{\partial \mathbf{q}} = \mathbf{F}_{\mathbf{q}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

This is an $m \times n$ matrix!

- Example 2.5.2:

$$\mathbf{q} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad \mathbf{r}^P = \begin{bmatrix} \cos \theta_1 + l \cos(\theta_1 + \theta_2) \\ \sin \theta_1 + l \sin(\theta_1 + \theta_2) \end{bmatrix} \quad \mathbf{r}_{\mathbf{q}}^P = ?$$

Putting Things in Perspective



Only a matter of notation: Left and Right mean the same thing

- Let x , y , and ϕ be three generalized coordinates

- Define the function \mathbf{r} of x , y , and ϕ as

$$\mathbf{r}(x, y, \phi) = \begin{bmatrix} x + 2l\cos\phi \\ y - 2l\sin\phi \end{bmatrix}$$

- Compute the partial derivatives

$$\mathbf{r}_{x,y,\phi} = \begin{bmatrix} \frac{\partial \mathbf{r}}{\partial x} & \frac{\partial \mathbf{r}}{\partial y} & \frac{\partial \mathbf{r}}{\partial \phi} \end{bmatrix}$$

- Let x , y , and ϕ be three generalized coordinates, and define the array \mathbf{q}

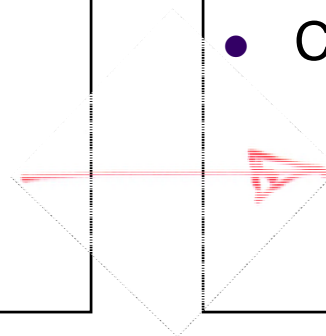
$$\mathbf{q} = \begin{bmatrix} x \\ y \\ \phi \end{bmatrix}$$

- Define the function \mathbf{r} of \mathbf{q} :

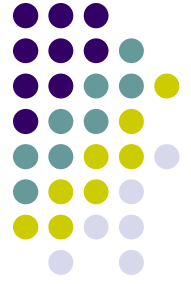
$$\mathbf{r}(\mathbf{q}) = \begin{bmatrix} x + 2l\cos\phi \\ y - 2l\sin\phi \end{bmatrix}$$

- Compute the partial derivative

$$\mathbf{r}_{\mathbf{q}} = \frac{\partial \mathbf{r}}{\partial \mathbf{q}}$$



Exercise



$$\mathbf{q} = \begin{bmatrix} x \\ y \\ \phi \end{bmatrix}$$

$$\mathbf{r}(\mathbf{q}) = \begin{bmatrix} x + 2l \cos \phi \\ y - 2l \sin \phi \end{bmatrix}$$

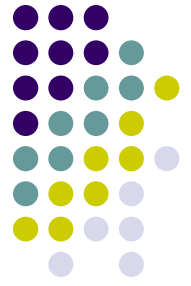
$$\mathbf{r}_{\mathbf{q}} = \frac{\partial \mathbf{r}}{\partial \mathbf{q}} = ?$$

$$\mathbf{r}_{\mathbf{q}} = \frac{\partial \mathbf{r}}{\partial \mathbf{q}} = \left[\frac{\partial \mathbf{r}}{\partial q_1} \quad \frac{\partial \mathbf{r}}{\partial q_2} \quad \frac{\partial \mathbf{r}}{\partial q_3} \right] = \left[\frac{\partial \mathbf{r}}{\partial x} \quad \frac{\partial \mathbf{r}}{\partial y} \quad \frac{\partial \mathbf{r}}{\partial \phi} \right]$$

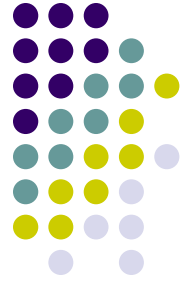


$$\mathbf{r}_{\mathbf{q}} = \begin{bmatrix} 1 & 0 & -2l \sin \varphi \\ 0 & 1 & -2l \cos \varphi \end{bmatrix}$$

Partial Derivatives: Good to Remember...

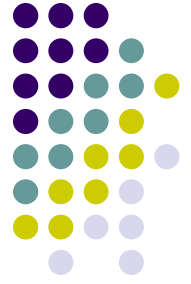


- Most general case: you start with “ m ” functions stacked up in a vector; each function depends on a set of “ n ” variables
- You end with an $m \times n$ matrix; each of its entries is a partial derivative
 - You start with a column vector of functions and end up with a matrix
- Taking a partial derivative leads to a *higher dimensional* quantity
 - Scalar Function – leads to row vector
 - Vector Function – leads to matrix
 - In ME451 we called this the “accordion rule”
- In this class, taking partial derivatives can lead to one of the following:
 - A row vector
 - A full blown matrix
 - In this class, if you see something else there is a mistake somewhere
- For partial derivative, so far we’ve only introduced definitions



Done with plain vanilla Partial Derivatives
... moving on to...
Partial Derivatives requiring the Chain Rule of Differentiation

Scenario 1: Scalar Function



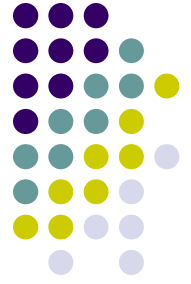
- f is a function of “ n ” variables: q_1, \dots, q_n
$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

- However, each of these variables q_i in turn depends on a set of “ k ” other variables x_1, \dots, x_k .

$$\mathbf{q} = \begin{bmatrix} q_1(x_1, \dots, x_k) \\ \dots \\ q_n(x_1, \dots, x_k) \end{bmatrix} : \mathbb{R}^k \rightarrow \mathbb{R}^n$$

- The composition of f and \mathbf{q} leads to a new function $\phi(\mathbf{x})$:

$$\phi(\mathbf{x}) = f \circ \mathbf{q} = f(\mathbf{q}(\mathbf{x})) : \mathbb{R}^k \rightarrow \mathbb{R}$$



Chain Rule for a Scalar Function

- The question: how do you compute $\phi_{\mathbf{x}}$?
 - Using our notation:

$$\phi = f \circ \mathbf{q} = f(\mathbf{q}(\mathbf{x})) \quad \Rightarrow \quad \phi_{\mathbf{x}} = \frac{\partial \phi}{\partial \mathbf{x}} = ??$$

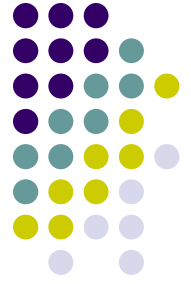
- Theorem: Chain rule of differentiation for scalar function

$$\phi_{\mathbf{x}} = \frac{\partial \phi}{\partial \mathbf{x}} = \frac{\partial f}{\partial \mathbf{q}} \cdot \frac{\partial \mathbf{q}}{\partial \mathbf{x}} = f_{\mathbf{q}} \cdot \mathbf{q}_{\mathbf{x}}$$

(Elementary calculus result)

[AO]

Example



Assume that $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and a function ϕ of \mathbf{y} is defined as: $\phi(\mathbf{y}) = 3y_1^2 + \sin y_2$.

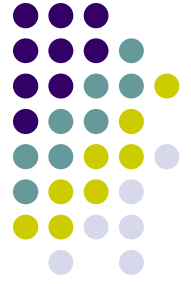
In turn, \mathbf{y} depends on a variable $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ as follows:

$$\mathbf{y} = \mathbf{y}(\mathbf{x}) = \begin{bmatrix} y_1(\mathbf{x}) \\ y_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} 2x_1 + \log x_2 + \sqrt{x_3} \\ (x_1 - x_2)^2 \end{bmatrix}$$

Now, since ϕ depends on \mathbf{y} and \mathbf{y} depends on \mathbf{x} , it means that ϕ depends on \mathbf{x} . Find the partial derivative of ϕ with respect to \mathbf{x} , that is,

$$\phi_{\mathbf{x}} = \frac{\partial \phi}{\partial \mathbf{x}} = \left[\frac{\partial \phi}{\partial x_1} \quad \frac{\partial \phi}{\partial x_2} \quad \frac{\partial \phi}{\partial x_3} \right] = ?$$

Scenario 2: Vector Function



- \mathbf{F} is a vector function of “ n ” variables: q_1, \dots, q_n

$$\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

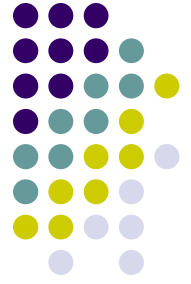
- However, each of these variables q_i in turn depends on a set of “ k ” other variables x_1, \dots, x_k .

$$\mathbf{q} = \begin{bmatrix} q_1(x_1, \dots, x_k) \\ \dots \\ q_n(x_1, \dots, x_k) \end{bmatrix} : \mathbb{R}^k \rightarrow \mathbb{R}^n$$

- The composition of \mathbf{F} and \mathbf{q} leads to a new function $\Phi(\mathbf{x})$:

$$\Phi(\mathbf{x}) = \mathbf{F} \circ \mathbf{q} = \mathbf{F}(\mathbf{q}(\mathbf{x})) : \mathbb{R}^k \rightarrow \mathbb{R}^m$$

Chain Rule for a Vector Function



- How do you compute the partial derivative of Φ ?

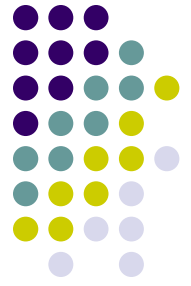
$$\Phi : \mathbb{R}^k \rightarrow \mathbb{R}^m$$

$$\Phi = \Phi(\mathbf{q}(\mathbf{x})) \quad \Rightarrow \quad \Phi_{\mathbf{x}} = \frac{\partial \Phi}{\partial \mathbf{x}} = ??$$

- Theorem: Chain rule of differentiation for vector functions

$$\Phi_{\mathbf{x}} = \frac{\partial \Phi}{\partial \mathbf{x}} = \frac{\partial \Phi}{\partial \mathbf{q}} \cdot \frac{\partial \mathbf{q}}{\partial \mathbf{x}}$$

Important Rule + Quick Examples



Important rule. Regarding how you can/should take partial derivatives:

If you need to take a partial derivative with respect to \mathbf{q} of a *quantity* that depends on \mathbf{q} , then \mathbf{q} should show up as the rightmost term in the *quantity*.

Example 1, good case: You can take the partial derivative of $\mathbf{B}\mathbf{q}$, where \mathbf{B} is a matrix that doesn't depend on \mathbf{q} . The result is:

Example 2, bad case: You can't take the partial derivative of $\mathbf{q}^T \mathbf{p}$ since \mathbf{q} is not the rightmost quantity; instead, \mathbf{p} is.

Example 3, good case: You can take the partial derivative of $\mathbf{p}^T \mathbf{q}$ since \mathbf{q} is the rightmost quantity. The result is:

Example 4, bad case: You can't take the partial derivative of $\mathbf{q}^T \mathbf{B}\mathbf{p}$ since \mathbf{q} is not the rightmost quantity; instead, \mathbf{p} is.

Example 5, good case: You can take the partial derivative of $\mathbf{p}^T \mathbf{B}^T \mathbf{q}$ since \mathbf{q} is the rightmost quantity. The result is:

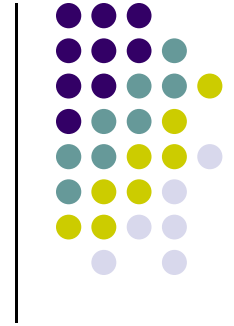
When Taking a Partial Derivative



- Understand with respect to what you are taking the partial derivative
 - Figure out its dimension
- Investigate the quantity that you want to take the partial derivative of
 - Figure out its dimension
 - Figure out what variables it depends on
- Remember the “rightmost only” rule described on the previous slide

[AO]

Example



Assume that $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and a function \mathbf{f} of \mathbf{y} is defined as: $\mathbf{f}(\mathbf{y}) = \begin{bmatrix} 2y_1 + y_2^2 \\ y_1y_2 \end{bmatrix}$.

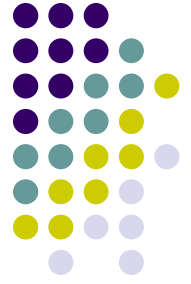
In turn, \mathbf{y} depends on a variable $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ as follows:

$$\mathbf{y} = \mathbf{y}(\mathbf{x}) = \begin{bmatrix} y_1(\mathbf{x}) \\ y_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} x_1x_2 \\ x_1^2 - x_2 \end{bmatrix}$$

Now, since \mathbf{f} depends on \mathbf{y} and \mathbf{y} depends on \mathbf{x} , it means that \mathbf{f} depends on \mathbf{x} . Find the partial derivative of \mathbf{f} with respect to \mathbf{x} , that is,

$$\mathbf{f}_{\mathbf{x}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \left[\frac{\partial \mathbf{f}}{\partial x_1} \quad \frac{\partial \mathbf{f}}{\partial x_2} \right] = ?$$

Scenario 3: Function of Two Vectors



- \mathbf{F} is a vector function of 2 vector variables \mathbf{q} and \mathbf{p} :

$$\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

- Both \mathbf{q} and \mathbf{p} in turn depend on a set of “ k ” other variables $\mathbf{x}=[x_1, \dots, x_k]^T$:

$$\mathbf{q} = \mathbf{q}(x_1, \dots, x_k) : \mathbb{R}^k \rightarrow \mathbb{R}^{n_1}$$

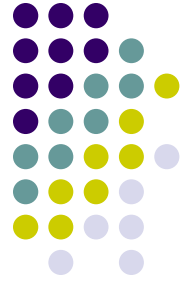
$$\mathbf{p} = \mathbf{p}(x_1, \dots, x_k) : \mathbb{R}^k \rightarrow \mathbb{R}^{n_2}$$

$$n = n_1 + n_2$$

- A new function $\Phi(\mathbf{x})$ is defined as:

$$\Phi(\mathbf{x}) = \mathbf{F}(\mathbf{q}(\mathbf{x}), \mathbf{p}(\mathbf{x})) : \mathbb{R}^k \rightarrow \mathbb{R}^m$$

The Chain Rule



- How do you compute the partial derivative of Φ with respect to \mathbf{x} ?

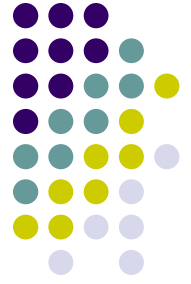
$$\Phi : \mathbb{R}^k \rightarrow \mathbb{R}^m$$

$$\Phi = \Phi(\mathbf{x}) \quad \Rightarrow \quad \Phi_{\mathbf{x}} = \frac{\partial \Phi}{\partial \mathbf{x}} = ??$$

- Theorem: Chain rule for function of two vectors

$$\Phi_{\mathbf{x}} = \frac{\partial \Phi}{\partial \mathbf{x}} = \frac{\partial \mathbf{F}}{\partial \mathbf{q}} \cdot \frac{\partial \mathbf{q}}{\partial \mathbf{x}} + \frac{\partial \mathbf{F}}{\partial \mathbf{p}} \cdot \frac{\partial \mathbf{p}}{\partial \mathbf{x}} = \mathbf{F}_{\mathbf{q}} \cdot \mathbf{q}_{\mathbf{x}} + \mathbf{F}_{\mathbf{p}} \cdot \mathbf{p}_{\mathbf{x}}$$

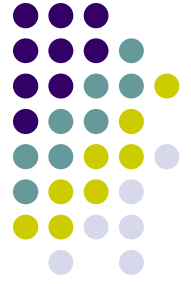
Example



Assume that $\mathbf{q} = \mathbf{q}(\mathbf{x}) \in \mathbb{R}^3$, and $\mathbf{p} = \mathbf{p}(\mathbf{x}) \in \mathbb{R}^3$. Show that:

$$\frac{\partial(\mathbf{q}^T \mathbf{p})}{\partial \mathbf{x}} = \mathbf{q}^T \mathbf{p}_{\mathbf{x}} + \mathbf{p}^T \mathbf{q}_{\mathbf{x}}$$

Scenario 4: Time Derivatives



- On the previous slides we talked about functions f of y , while y in turn depended on yet another variable x
- The relevant case is when the variable x is actually time, t
 - This scenario is super common in 751:
 - You have a function that depends on the generalized coordinates \mathbf{q} , and in turn the generalized coordinates are functions of time (they change in time, since we are talking about kinematics/dynamics here...)
 - Case 1: scalar function that depends on an array of m generalized coordinates that in turn depend on time

$$\Phi = \Phi(\mathbf{q}(t)) \in \mathbb{R}$$

- Case 2: vector function (of dimension n) that depends on an array of m generalized coordinates that in turn depend on time

$$\Phi = \Phi(\mathbf{q}(t)) \in \mathbb{R}^n$$



A Special Case: Time Derivatives (Cntd)

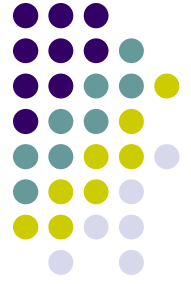
- Quantities of interest: the time derivative of Φ and Φ
- Apply the chain rule, the scalar function Φ case first:

$$\dot{\Phi} = \frac{d\Phi}{dt} = \frac{d\Phi(\mathbf{q}(t))}{dt} = \frac{\partial\Phi}{\partial\mathbf{q}} \cdot \frac{d\mathbf{q}}{dt} = \Phi_{\mathbf{q}}\dot{\mathbf{q}} \in \mathbb{R}$$

- For the vector function case, applying the chain rule leads to the same formula, only the size of the result is different...

$$\dot{\Phi} = \frac{d\Phi}{dt} = \frac{d\Phi(\mathbf{q}(t))}{dt} = \frac{\partial\Phi}{\partial\mathbf{q}} \cdot \frac{d\mathbf{q}}{dt} = \Phi_{\mathbf{q}}\dot{\mathbf{q}} \in \mathbb{R}^n$$

Example, Scalar Function Φ



- Assume a set of generalized coordinates is defined through array \mathbf{q} . Also, a scalar function Φ of \mathbf{q} is provided:

$$\mathbf{q}(t) = \begin{bmatrix} x(t) \\ y(t) \\ \theta(t) \end{bmatrix}$$

$$\Phi(\mathbf{q}) = 3x(t) + 2L \sin \theta(t)$$

- Find time derivative of Φ

$$\dot{\Phi} = ?$$

Example, Vector Function Φ



- Assume a set of generalized coordinates is defined through array \mathbf{q} . Also, a vector function Φ of \mathbf{q} is provided:

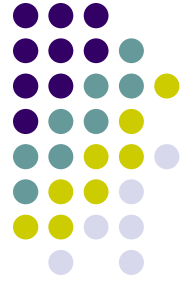
$$\mathbf{q}(t) = \begin{bmatrix} x(t) \\ y(t) \\ \theta(t) \end{bmatrix}$$

$$\Phi(\mathbf{q}) = \begin{bmatrix} 3x(t) & + & 2L \sin \theta(t) \\ y(t) & - & 2L \cos \theta(t) \end{bmatrix}$$

- Find time derivative of Φ

$$\dot{\Phi} = ?$$

Useful Formulas



- A couple of useful formulas, some of them you had to derive as part of the HW

$$\frac{\partial(\mathbf{g}^T \mathbf{p})}{\partial \mathbf{q}} = \mathbf{g}^T \mathbf{p}_{\mathbf{q}} + \mathbf{p}^T \mathbf{g}_{\mathbf{q}}$$

$$\frac{\partial}{\partial \mathbf{q}} (\mathbf{C}\mathbf{q}) = \mathbf{C}$$

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^T \mathbf{C}\mathbf{y}) = \mathbf{y}^T \mathbf{C}^T$$

$$\frac{d}{dt} (\mathbf{p}^T \mathbf{C}\mathbf{q}) = \dot{\mathbf{p}}^T \mathbf{C}\mathbf{q} + \mathbf{p}^T \mathbf{C}\dot{\mathbf{q}}$$

Assumptions:

$$\mathbf{g} = \mathbf{g}(\mathbf{q})$$

$$\mathbf{p} = \mathbf{p}(\mathbf{q})$$

\mathbf{C} - constant matrix

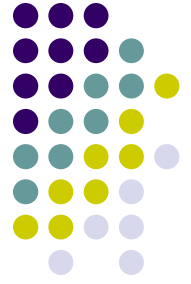
\mathbf{y} doesn't depend on \mathbf{x}

The dimensions of the vectors and matrix above such that all the operations listed can be carried out.

Example

- Derive the last equality on previous slide
- Can you expand that equation further?

$$\frac{d}{dt}(\mathbf{p}^T \mathbf{C} \mathbf{q}) = \dot{\mathbf{p}}^T \mathbf{C} \mathbf{q} + \mathbf{p}^T \mathbf{C} \dot{\mathbf{q}}$$



Assumptions:
 $\mathbf{p} = \mathbf{p}(\mathbf{q})$
 \mathbf{C} - constant matrix