2. Numerical Solution of Inverse Kinematics

Game Engineering
Prof. HyeongYeop Kang
siamiz@khu.ac.kr
IIIXR LAB

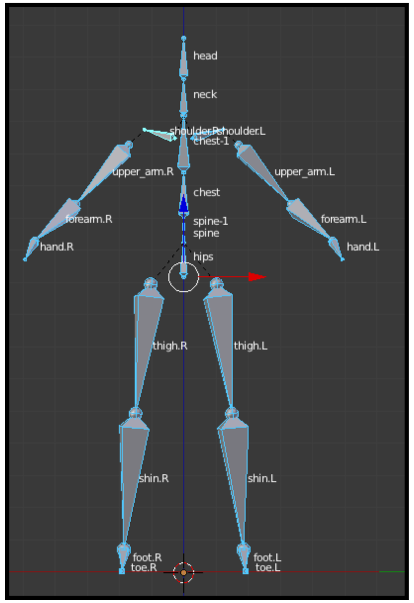
Introduction

Introduction

- Let the complete joint configuration of the multibody be specified by the scalars $\theta_1, \ldots, \theta_n$, assuming that there are n joints and each θ_j value is called a **joint angle**, where θ_j is the angle in the plane of rotation assuming we also have knowledge of the rotation axis.

- Certain points on the links are identified as end effectors.

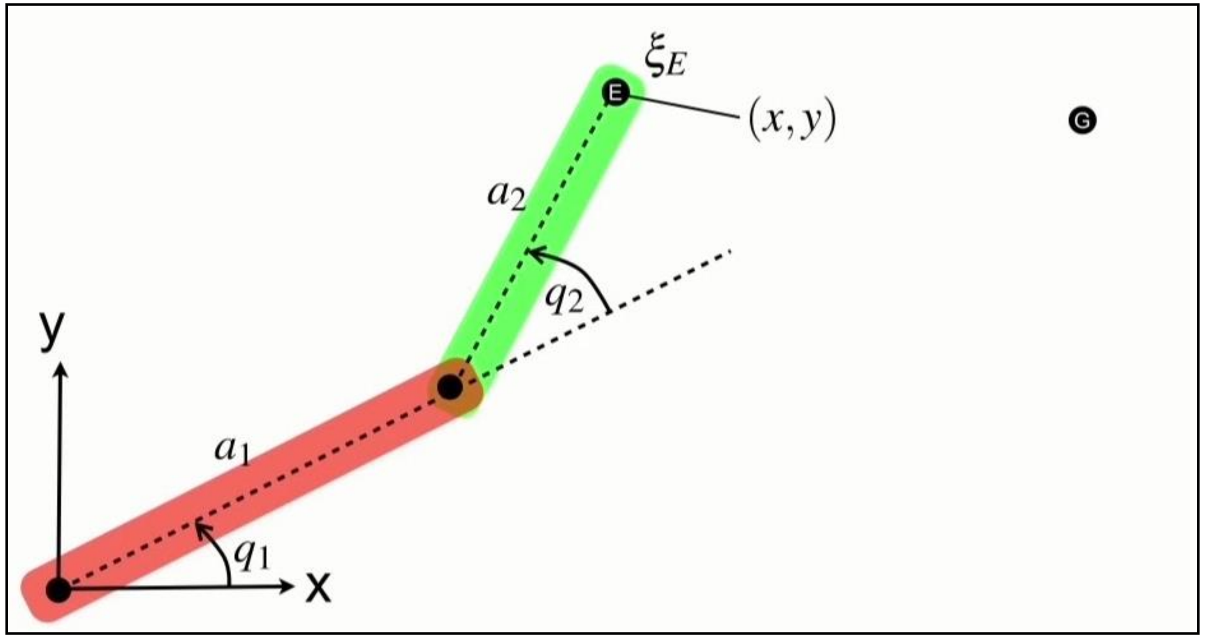
- To solve the IK problem, the joint angles must be settled so that the resulting configuration of the multibody places each end effector as close as possible to its target position.
    - If there are k end effectors, let their positions be denoted as $s_1, \ldots, s_k$ relative to a fixed origin.
    - The end effector positions are defined by a column vector $(s_1, s_2, \ldots, s_k)^T$ and can be written as $\vec{s}$.
    - The target (goal) positions are also defined by a column vector $(t_1, t_2, \ldots, t_k)^T$, where t_i is the target position for the i-th end effector.
    - The joint angles are also written as a column vector $\theta = (\theta_1, \ldots, \theta_n)^T$.
    - The desired change in position of the i-th end effector can be written as $e_i = t_i - s_i$.
    - In Forward Kinematics (FK), the end effector positions are functions of the joint angles: $\vec{s} = f(\theta)$.
    - In Inverse Kinematics (IK), the goal is to find a vector given desired configuration $\vec{s}_d$: $\theta = f^{-1}(\vec{s}_d)$.



Introduction

Introduction

- To solve IK problem, iterative methods to approximate a good solution to the problem seems required.
  - The function f() is a highly non linear operator.
  - There is an instance where a solution to the IK problem does not exist due to an unreachable target.
  - There is an instance where the best solution is not unique.
- To understand IK problem, you must understand kinetic angular variables first.



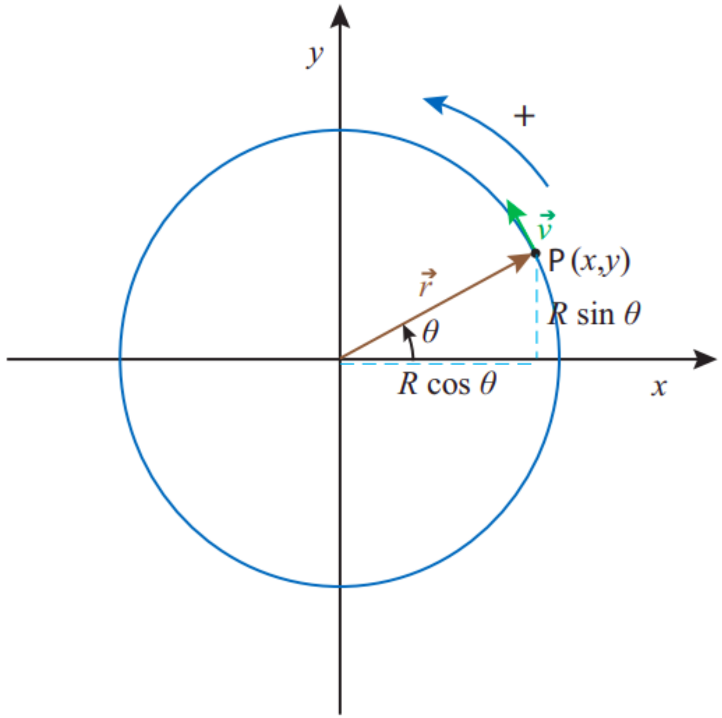
Kinetic Angular Variables

Kinetic Angular Variables[3]

Consider a particle moving on a circle. It can be represented by Cartesian coordinates (x, y). However, a value of x and y does not immediately tell us how far the particle has traveled along the circle itself.

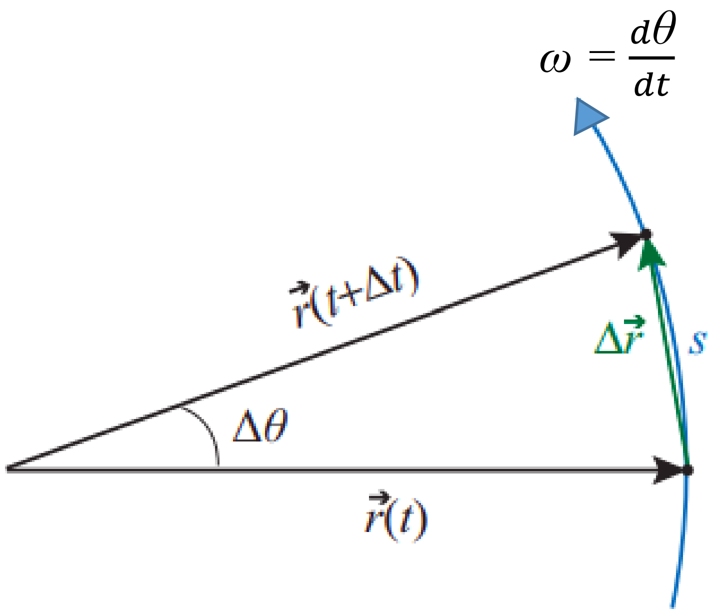
Instead, the most convenient way to describe the motion of the particle is to give the angle θ that the position vector makes with some reference axis at any given time.

If we choose the x axis as the reference, then the conversion from a description based on the radius R and the angle θ to a description in terms of x and y is simply $x = R\cos\theta$ and $y = R\sin\theta$.
(Note that the standard convention is that θ grows in the counterclockwise direction from the reference axis.)



Kinetic Angular Variables

Kinetic Angular Variables[3]



Information on the direction of motion at any given time is captured by the concept of the **angular velocity**, which we represent by the symbol ω and define in a manner analogous to the way we defined the ordinary velocity: if $\Delta\theta = \theta(t+\Delta t) - \theta(t)$ is the angular displacement over a time Δt, then $\omega = \lim_{\Delta t \to 0} \frac{\Delta\theta}{\Delta t} = \frac{d\theta}{dt}$.

For motion with constant angular velocity, we clearly will have

$$\theta(t) = \theta_i + \omega(t - t_i) \text{ or } \Delta\theta = \omega\Delta t$$

where ω changes with time, we can introduce an **angular acceleration** $\alpha = \lim_{\Delta t \to 0} \frac{\Delta\omega}{\Delta t} = \frac{d\omega}{dt}$.

Then for motion with constant angular acceleration we have the formulas:

$$\omega(t) = \omega_i + \alpha(t - t_i) \text{ or } \Delta\omega = \alpha\Delta t$$

$$\theta(t) = \theta_i + \omega(t - t_i) + \frac{1}{2}\alpha(t - t_i)^2 \text{ or } \Delta\theta = \omega_i\Delta t + \frac{1}{2}\alpha(\Delta t)^2$$

If we expressed θ in radians then the length of the arc corresponding to $\Delta\theta$ would be: $s = R|\Delta\theta|$.

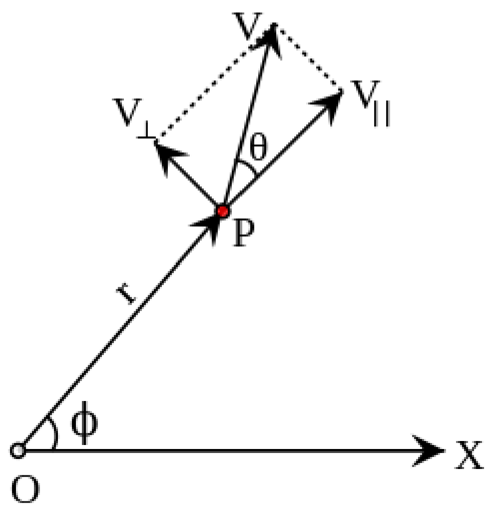
For very small angular displacements, $s \cong |\Delta\vec{r}|$. Dividing by Δt, using $s = R|\Delta\theta|$ and taking the $\Delta t \to 0$, we get the relationship between the angular velocity and the instantaneous speed v: $|\vec{v}| = |\lim_{\Delta t \to 0} \frac{s}{\Delta t}| = R|\omega|$.

Kinetic Angular Variables

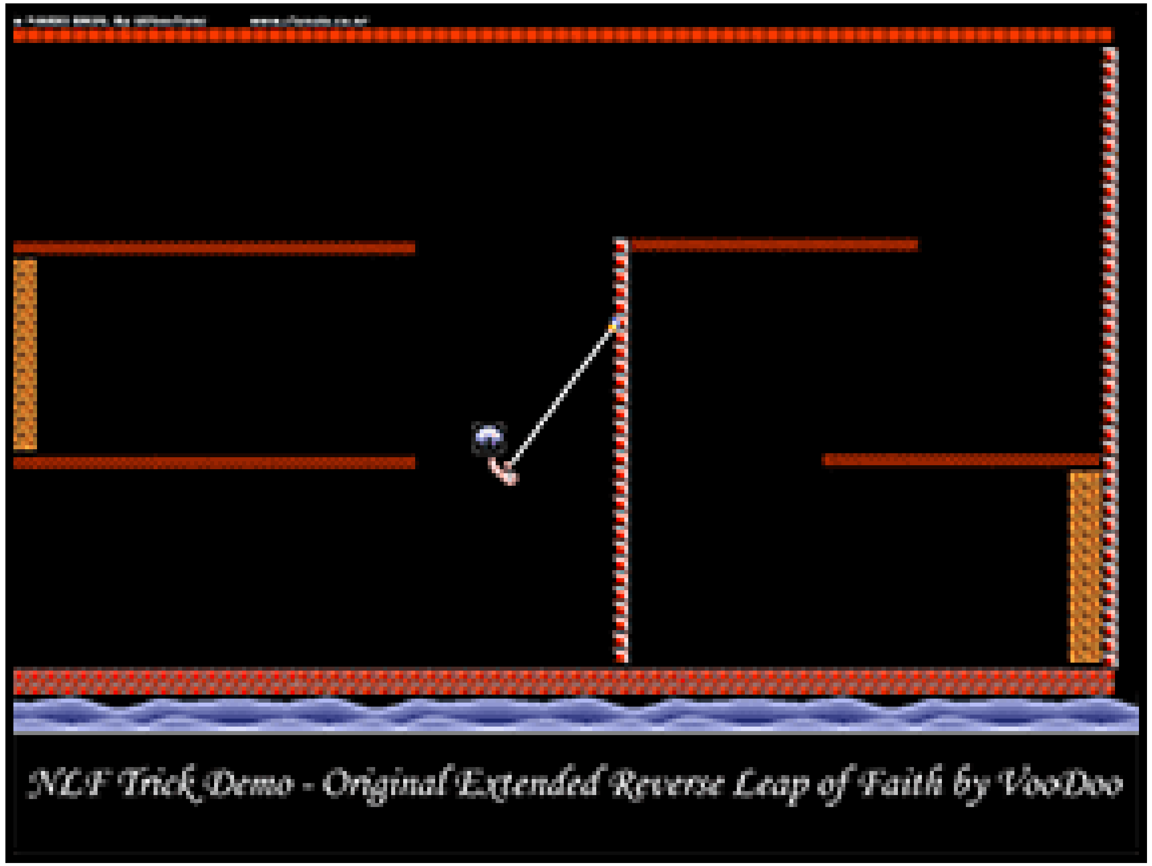
Angular velocity in two dimensions[4]

The diagram presented on the right shows r (position vector from the origin O to a particle P), with its polar coordinates (r,ϕ). The particle has linear velocity splitting as $v = v_{\parallel} + v_{\perp}$, with the radial component $v_{\parallel}$ and the tangential component $v_{\perp}$. Then the angular velocity can be computed from the cross-radial velocity as: $\omega = \frac{d\phi}{dt} = \frac{v_{\perp}}{r} = \frac{v\sin(\theta)}{r}$

(from $\vec{v}_{\perp} = r\omega$)



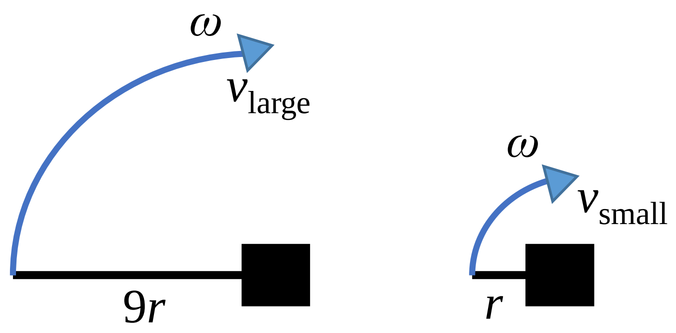
Angular Variables in Games



Angular Velocity - Practice

Q. A catapult of pole length $9r$ throws a watermelon from one end of a pole which rotates about an axis at the other end with angular speed ω. A smaller catapult of pole length r throws another watermelon from one end of a pole which rotates about an axis at the other end with the same angular speed ω as the first catapult.

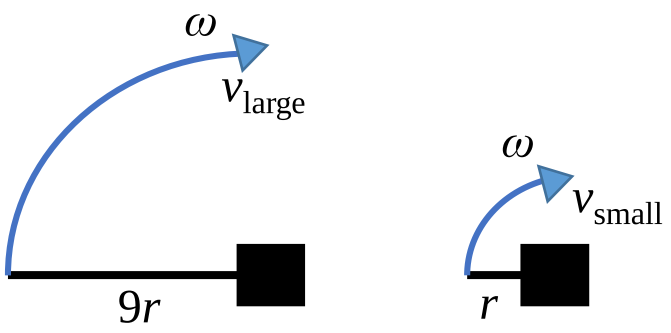
How does the linear speed of the larger catapult's watermelon v_{large} compare with the smaller catapult's watermelon v_{small}?



Angular Velocity – Practice Solution

Q. A catapult of pole length 9*r* throws a watermelon from one end of a pole which rotates about an axis at the other end with angular speed ω. A smaller catapult of pole length *r* throws another watermelon from one end of a pole which rotates about an axis at the other end with the same angular speed ω as the first catapult.

How does the linear speed of the larger catapult's watermelon v_{large} compare with the smaller catapult's watermelon v_{small}?



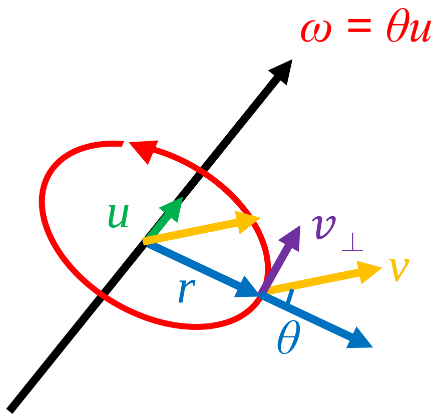
A. $v_{\text{large}} = 9r\omega$, $v_{\text{small}} = r\omega$

Kinetic Angular Variables

Angular velocity in three dimensions[4]

In three-dimensional space, rotation angle can be represented by axis-angle representation. It is parameterized by two quantities.

- A unit vector indicating the direction of an axis of rotation: u
- An angle describing the magnitude of the rotation about the axis: θ

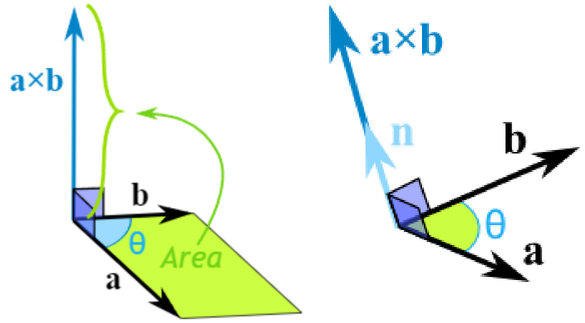


Then, the orbital angular velocity vector as

$$\omega = \omega_{2\mathrm{d}} u = \frac{d\theta}{dt} u = \frac{v\sin(\theta)}{r} u,$$

where ω_{2d} is 2D angular velocity defined in the rotational plane. Since the θ is the angle between r and v, ω can be written as:

$$\omega = \frac{v\sin(\theta)}{r} u = \frac{r v\sin(\theta) u}{r r} = \frac{r \times v}{r^2}$$



Note that a × b = |a| |b| sin(θ)n. Then, we can recover the tangential velocity as: $\omega \times r = |\frac{v\sin(\theta)}{r}$u| |r| sin(90°) $\hat{v}_{\perp} = v_{\perp}$

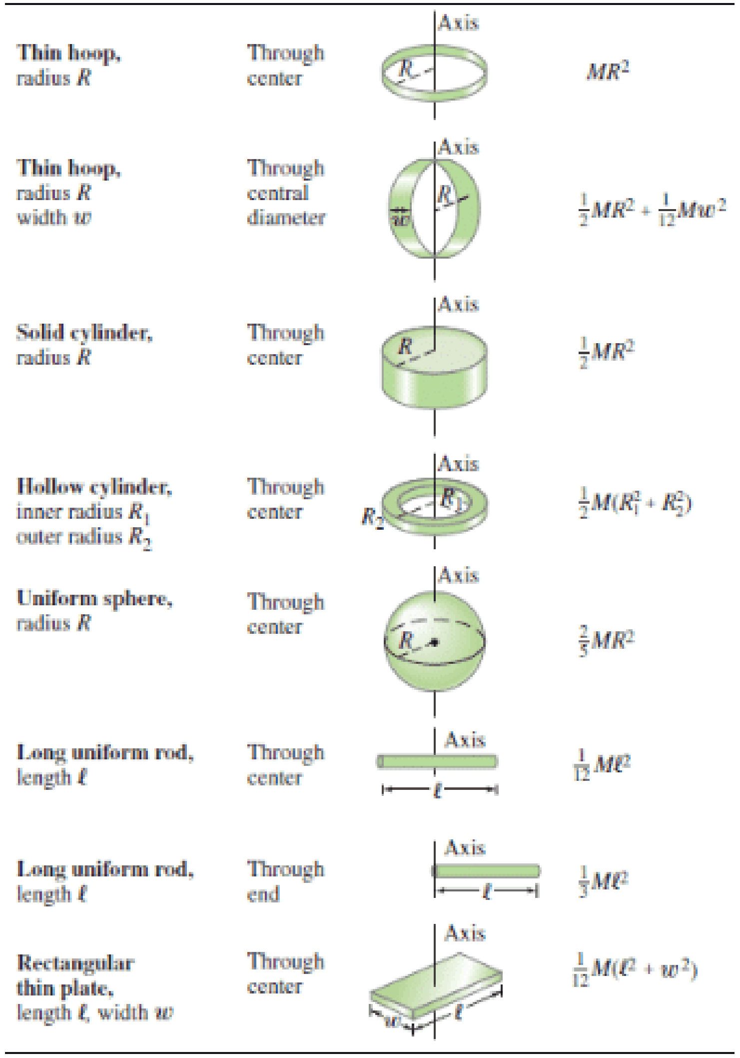
Angular Momentum

Angular momentum and moment of inertia[5]

Angular momentum is a vector quantity that represents the product of a body's rotational inertia and rotational velocity about a particular axis. However, if the particle's trajectory lies in a single plane, it is sufficient to discard the vector nature of angular momentum, and treat it as a scalar (more precisely, a pseudoscalar).

Angular momentum can be considered a rotational analog of linear momentum. Thus, where linear momentum p is proportional to mass m and linear speed v, $p = mv$, angular momentum L is proportional to moment of inertia I and angular speed ω measured in radians per second, $L = I\omega$.

The moment of inertia characterize the resistance of a body to changes in its motion. It depends on how mass is distributed around an axis of rotation, and will vary depending on the chose axis. For a point-like mass, the moment of inertia about some axis is given by mr^2.



Basics

Gradient[gra]

- The gradient of a scalar-valued multivariable function $f(x,y, \ldots)$, denoted ∇f, packages all its partial derivate information into a vector:

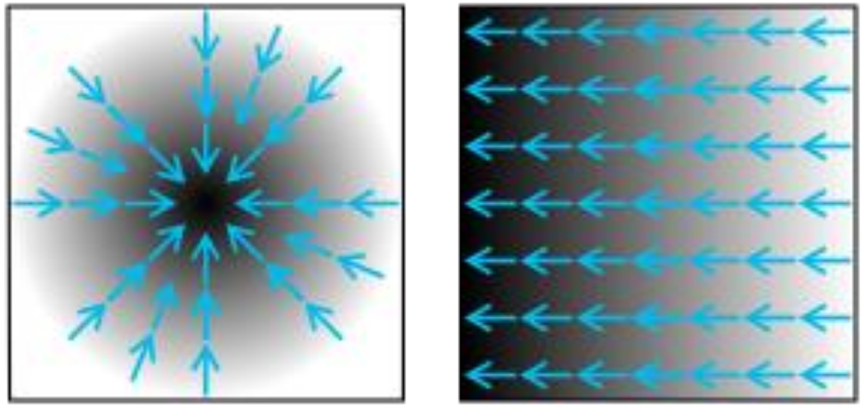
$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \vdots \end{bmatrix}$$

- For example, if $f(x, y) = 2x^2 - 4xy$, ∇f can be written as $\begin{bmatrix} 4x - 4y \\ -4x \end{bmatrix}$.

- If you imagine standing at a point $(x_0, y_0, \ldots)$ in the input space of f, the vector $\nabla f\,(x_0, y_0, \ldots)$ tells you which direction you should travel to increase the value of f most rapidly.

- ∇f can be visualized with a vector field called gradient field of f.

- In the figure shown right, the gradient is represented by the blue arrows that denote the direction of the greatest change of a scalar function. Note that the values of a function are represented in greyscale and increase in value from white to dark[gravis].



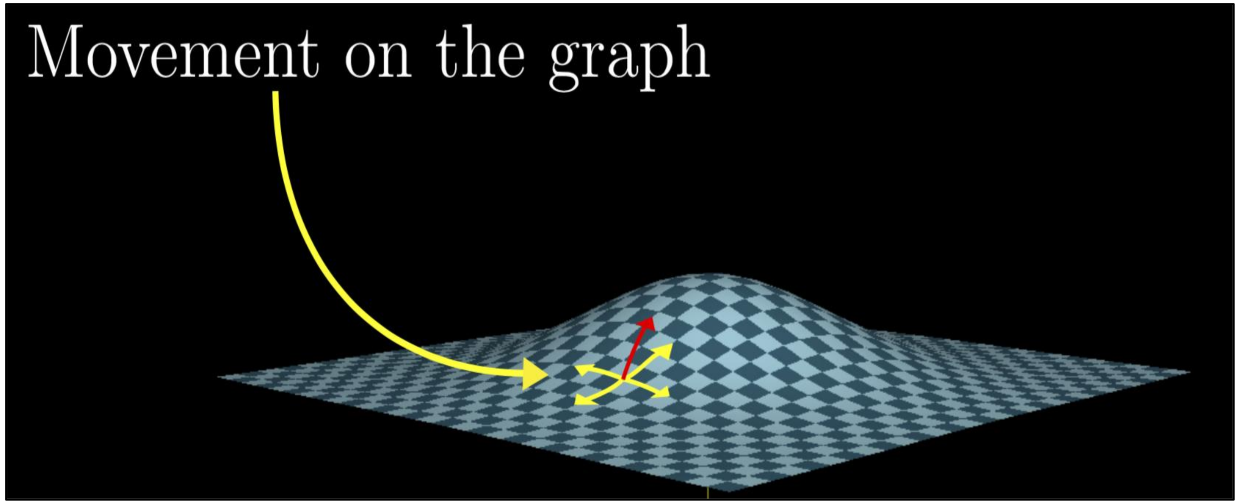
Basics

Interpreting the gradient[gra]

- Let's think about the case where the input of f is two-dimensional. The gradient turns each input point (x_0, y_0) into the vector:

$$\nabla f(x_o, y_0) = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix}$$

- Then, the vector tells us about the behavior of the function around the point (x_0, y_0):
  - Let's think about the function f that represents a hilly terrain surface. If you are standing on the point (x_0, y_0), the slope of the hill depends on which direction you walk; if you step straight in the positive y-direction, the slope is $\frac{\partial f}{\partial y}$.
  - If you walk in the direction of the gradient, you walk up the hill in the fastest direction.



Basics

Hessian[hesi]

- The Hessian matrix or Hessian is a square matrix of second-order partial derivatives of a scalar-valued function or scalar field.
  - It describes the local curvature of a function of many variables. $(\mathbf{H}_f)_{i,j} = \frac{\partial^2 f}{\partial x_i \, \partial x_j}.$

- Suppose f: $R^n \to R$ is a function taking as input a vector x and outputting a scalar $f(x) \in R$.
  - If all second partial derivatives of f exist and are continuous over the domain of the function, then the Hessian matrix H of f is a square $n \times n$ matrix.
  - The Hessian matrix is a symmetric matrix since the hypothesis of continuity of the second derivatives implies that the order of differentiable does not matter ($\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$).
  - The determinant of the Hessian matrix is called the Hessian determinant.
  - The Hessian matrix of a function f is the Jacobian matrix of the gradient of the function f; that is $H(f(x)) = J(\nabla f(x))$ – *See next page.*

$$\mathbf{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \, \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \, \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \, \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \, \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \, \partial x_1} & \frac{\partial^2 f}{\partial x_n \, \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Basics

Jacobian and Hessian

- The Hessian matrix of a function f is the Jacobian matrix of the gradient of the function f; that is H(f(x)) = J($\triangledown f(x)$).

- Let's compute H(f(x)) when $f(x,y) = x^3 + 2x^2y + 3xy + 4y^3$
  - Let's start to compute the first order partial derivatives of the function:

$$\frac{\partial f}{\partial x} = 3x^2 + 4xy + 3y, \quad \frac{\partial f}{\partial y} = 12y^2 + 2x^2 + 3x$$

  - Then we can compute the second-order partial derivatives of the function:

$$\frac{\partial^2 f}{\partial x^2} = 6x + 4y, \quad \frac{\partial^2 f}{\partial y^2} = 24y, \quad \frac{\partial^2 f}{\partial x \partial y} = 4x + 3, \quad \frac{\partial^2 f}{\partial y \partial x} = 4x + 3$$

  - Now, we can find the Hessian matrix using the formula for 2×2 matrices:

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 6x + 4y & 4x + 3 \\ 4x + 3 & 24y \end{bmatrix}$$

Basics

Jacobian and Hessian

- Let's compute $J(\triangledown f(x))$ when $f(x,y) = x^3 + 2x^2y + 3xy + 4y^3$
  - Let's start to compute the gradient of the f with first order partial derivatives of the function:

$$\frac{\partial f}{\partial x} = 3x^2 + 4xy + 3y,\ \frac{\partial f}{\partial y} = 12y^2 + 2x^2 + 3x$$

$$\triangledown f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) \end{bmatrix} = \begin{bmatrix} 3x^2 + 4xy + 3y \\ 12y^2 + 2x^2 + 3x \end{bmatrix}$$

  - Recall that the Jacobian matrix of $y = f(x)$ comprising three equations (f_1, f_2, f_3) with three variables (x,y,z) is defined by

$$J(\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{bmatrix}$$

  - Then the Jacobian matrix of the gradient of the function, $J(\triangledown f(x))$, can be computed by

$$J(\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}) = \begin{bmatrix} \frac{\partial^2 f_1}{\partial^2 x} & \frac{\partial^2 f_1}{\partial x \partial y} \\ \frac{\partial^2 f_2}{\partial y \partial x} & \frac{\partial^2 f_2}{\partial^2 y} \end{bmatrix} = \begin{bmatrix} 6x + 4y & 4x + 3 \\ 4x + 3 & 24y \end{bmatrix}$$

  - Therefore, $H(f(x)) = J(\triangledown f(x))$.

Gradient & Hessian - Practice

Q. Find the gradient vector of $x^2y + 4ze^{x+y} = f(x, y, z)$ at point (3, -3, 2).

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \vdots \end{bmatrix}$$

Q. Let $F(x, y) = \ln(x)\ln(y)$. What is the Hessian of F?

- Q1. $\mathrm{H}(F) = \begin{bmatrix} ??? & ??? \\ ??? & ??? \end{bmatrix}$

$$\mathbf{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\,\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\,\partial x_1} & \frac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Gradient & Hessian – Practice Solution

Q. Find the gradient vector of $x^2y + 4ze^{x+y} = f(x, y, z)$ at point (3, -3, 2).

- $\nabla f = (2xy + 4ze^{x+y}, x^2 + 4ze^{x+y}, 4e^{x+y}) = (-10, 17, 4)$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \vdots \end{bmatrix}$$

Q. Let $F(x, y) = \ln(x)\ln(y)$. What is the Hessian of F?

- Q1. $\mathrm{H}(F) = \begin{bmatrix} ??? & ??? \\ ??? & ??? \end{bmatrix}$

$$\begin{bmatrix} \frac{-\ln(y)}{x^2} & \frac{1}{xy} \\ \frac{1}{xy} & \frac{-\ln(x)}{y^2} \end{bmatrix}$$

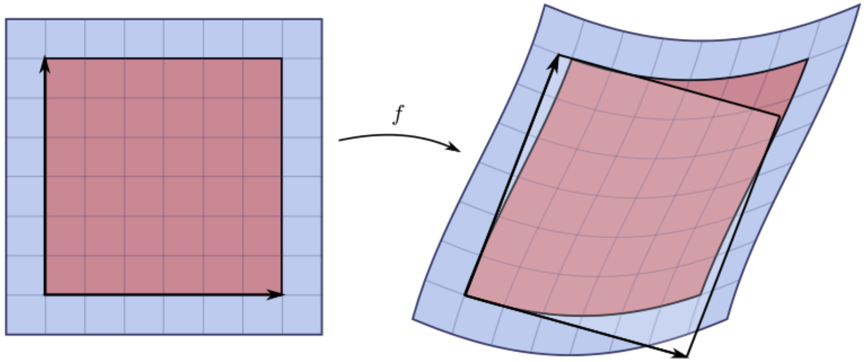
$$\mathbf{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\,\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\,\partial x_1} & \frac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

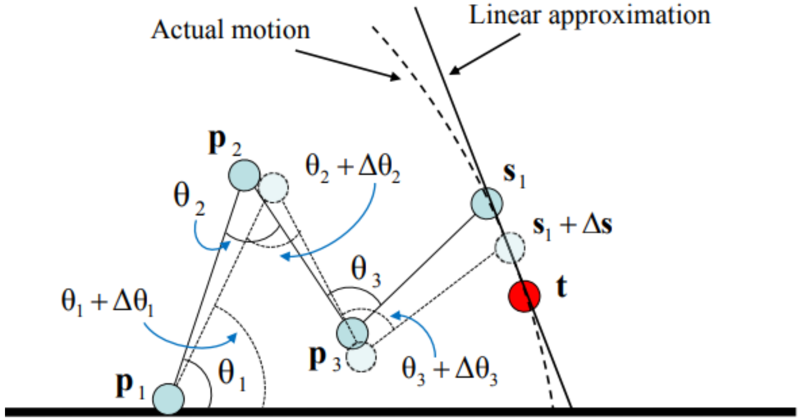
Jacobian Inverse Methods

Problem defined by Jacobian

- The Jacobian J is a matrix of partial derivatives of the entire chain system relative to the end effectors s.

- The Jacobian solutions are a *linear approximation* of the IK problem; they linearly model the end effectors' movements relative to instantaneous system changes in link translation and joint angle.





Jacobian Inverse Methods

Problem defined by Jacobian

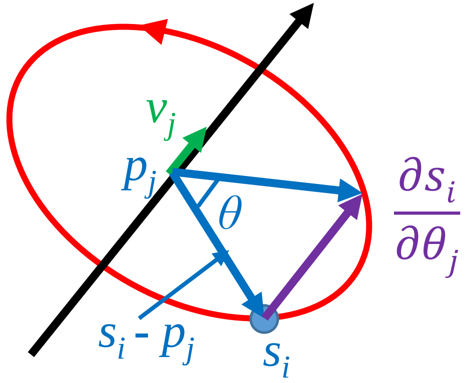
- The Jacobian matrix J is a function of the θ values and is defined by:

$$J(\theta)_{ij} = \left(\frac{\partial s_i}{\partial \theta_j}\right)_{ij} \quad (i = 1, \ldots, k \text{ and } j = 1, \ldots, n)$$

- If the j-th joint is a rotational joint with a single degree of freedom, the joint angle is a single scalar θ_j. Then, the Jacobian matrix entries for the j-th rotational joint can be calculated as follows[1]:

$$\omega_{2d} = \frac{\partial s_i}{\partial \theta_j} = v_j \times (s_i - p_j)$$

s_i: the position of the end effector
p_j: the position of the joint
v_j: the unit vector pointing along the current axis of rotation for the joint.
(assume that the angular velocity is only proportional to the s_i - p_j)



Jacobian Inverse Methods

IK Problem defined by Jacobian

- Now, $\vec{s}$ = f(θ) for forward dynamics can be written as : $\frac{d\vec{s}}{dt_t} = J(\theta)\frac{d\theta}{dt_t}$ where $J(\theta) = \frac{d\vec{s}}{d\theta}$ *and* t_t = *time.*
- The change in end effector positions caused by the update value $\Delta\theta$ can be estimated as $\Delta\vec{s} \approx J\Delta\theta$
- As our goal is to find the desired change in position of the effectors $\vec{e} = \vec{t} - \vec{s}$, $\Delta\vec{s}$ should be approximately equal to $\vec{e}$.
- Therefore, the FK problem can be expressed as $\vec{e} = J\Delta\theta$ and, consequently, the IK problem can be rewritten as $\Delta\theta = J^{-1}\vec{e}$.

Recall

The end effector positions are defined by a column vector $(s_1, s_2, ..., s_k)^T$ and can be written as $\vec{s}$.
The target (goal) positions are also defined by a column vector $(t_1, t_2, ..., t_k)^T$, where t_i is the target position for the i-th end effector.
The joint angles are also written as a column vector $\theta = (\theta_1, ..., \theta_n)^T$.
The desired change in position of the i-th end effector can be written as $e_i = t_i - s_i$.
In Forward Kinematics (FK), the end effector positions are functions of the joint angles: $\vec{s} = f(\theta)$.
In Inverse Kinematics (IK), the goal is to find a vector given desired configuration $\vec{s}_d$: $\theta = f^{-1}(\vec{s}_d)$.

Jacobian Inverse Methods

Jacobian Pseudo-inverse

- Unfortunately, in $\Delta\theta = J^{-1}\vec{e}$, the Jacobian J may not be square or invertible.

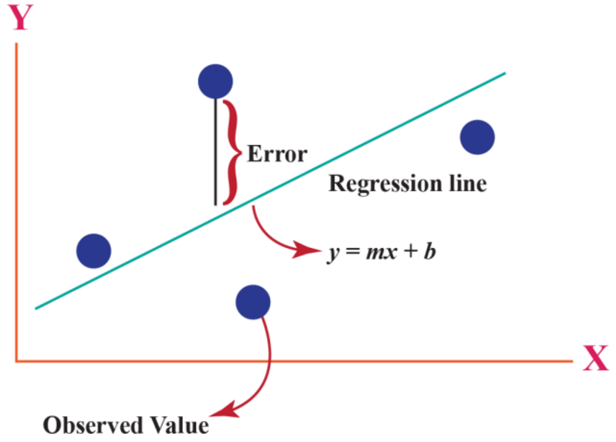
- Even if it is invertible, J may work poorly as it may be nearly singular (when no change in joint angle can achieve a desired change in chain end position.)

- To overcome these problems, the Jacobian Pseudo-inverse (also known as the Moore-Penrose inverse) of the Jacobian can be exploited.

- Let J^{+} is an $n \times m$ matrix called the pseudo-inverse of J. This pseudo-inverse gives the best possible solution to the equation $\vec{e} = J\Delta\theta$ in the least squares sense.

Least square?

The least squares method is a statistical procedure to find the best fit for a set of data points by minimizing the sum of the offsets or residuals of points from the plotted curve[2].



Jacobian Inverse Methods

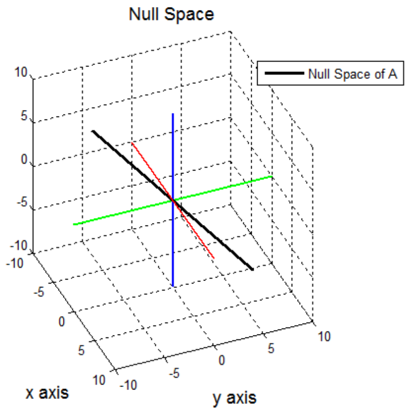
Jacobian Pseudo-inverse

- Pseudo-inverse has the property that the matrix $(I - J^{+}J)$ performs a projection onto the null space of J.
  - By using pseodu-inverse jacobian, $\Delta\theta$ of $\vec{e} = J\Delta\theta$ has unique solution when $\vec{e}$ is in the range of J and $\Delta\theta$ has the property that it minimizes the magnitude of the difference $J\Delta\theta - \vec{e}$ when $\vec{e}$ is not in the range of J.

- In one study[pseudoIK], the authors proposed a control policy by using a special form of the general solution of sets of linear equations: $J(I - J^{+}J)\varphi = 0$, for all vectors φ.
  - This means that we can set $\Delta\theta = J^{+}\vec{e} + (I - J^{+}J)\varphi$ (from $\Delta\theta = J^{-1}\vec{e}$) for any vector φ and still obtain a value for $\Delta\theta$ which minimize the value $J\Delta\theta - \vec{e}$.
  - Proof: $\Delta\theta = J^{+}\vec{e} + (I - J^{+}J)\varphi \rightarrow J\Delta\theta = JJ^{+}\vec{e} + (J - JJ^{+}J)\varphi = JJ^{+}\vec{e} = \vec{e}$

Null space of a matrix?

Null space is a concept in linear algebra which is used to identify the linear relationship among attributes. The null space of any matrix A consists of all the vectors B such that AB = 0. The null space of A always contains the zero vector, since A0 = 0.



Jacobian Inverse Methods

Jacobian Pseudo-inverse

- There are several approaches to obtain a J^+.
  - Analytical Jacobian Method (usually used for the simple robot problem)
  - Moore-Penrose Method
  - Jacobian Transpose Method
  - Damped Least Square Method
  - Singular Value Decomposition (SVD) Method

- These approaches will be handled in following lectures.

Reference

[1] Efficient computation of the jacobian for robot manipulators, 1984
[2] https://www.cuemath.com/data/least-squares/
[3] Motion on a Circle (Or Part of a Circle). (2020, November 6). University of Arkansas. https://phys.libretexts.org/@go/page/22249
[4] https://en.wikipedia.org/wiki/Angular_velocity
[5] https://en.wikipedia.org/wiki/Angular_momentum
[6] https://mathworld.wolfram.com/EigenDecomposition.html
[7] https://www.youtube.com/watch?v=PFDu9oVAE-g
[gra] https://www.khanacademy.org/math/multivariable-calculus/multivariable-derivatives/partial-derivative-and-gradient-articles/a/the-gradient
[gravis] https://en.wikipedia.org/wiki/Gradient
[hesi] https://en.wikipedia.org/wiki/Hessian_matrix