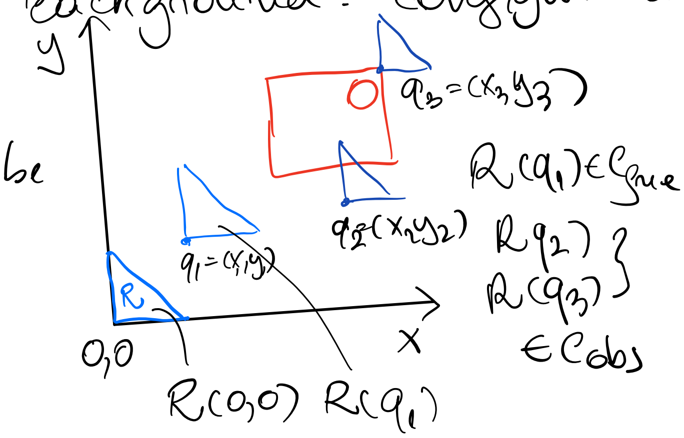
Trajectory Optimization

- Problem statement
- Multivariate calculus refresher
- Functional Gradient Descent
- CHOMP

Problem Statement

Background: configuration space



world

$R(q): C \to \mathcal{P}(W)$ $\mathbb{R}^2/\mathbb{R}^3$

$q \mapsto \{x \mid x \in W, x \text{ occupied by the robot}\}$

$q=(x,y)$ $R(q)=$

$q=(x,y,\theta)$ $R(q)=$

$q=(\theta_1,\theta_2)$ $R(q)=$

human arm?

Forward Kinematics: $\phi: C \to W$ (OR SE(2/3) - rotations)

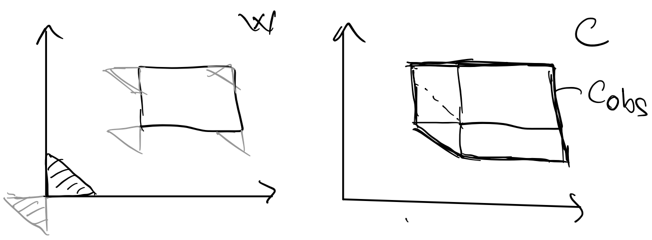
$\phi_{EE}: q \mapsto x$ — location of EE

$(\theta_1,\theta_2) \mapsto$ x

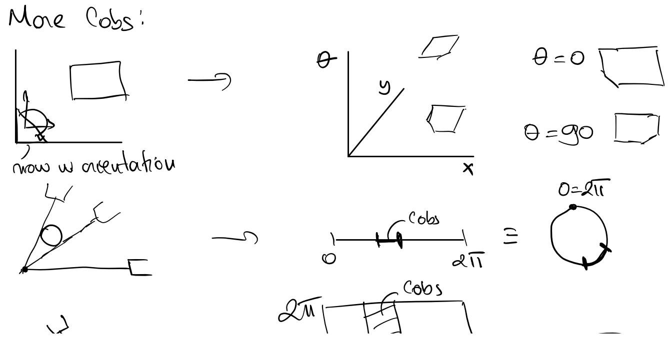
Inverse kinematics: IK: $W \to C$ / $\mathcal{P}(C)$

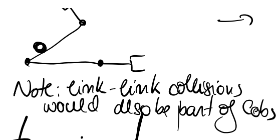
$IK(x) = \{q \mid \phi(q) = x\}$

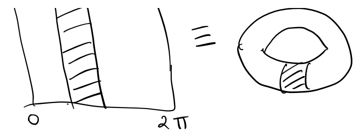
Obstacles $O \subset W$ $C_{obs} = \{q \in C \mid R(q) \cap O \neq \emptyset\}$ $C_{free} = C \setminus C_{obs}$



More Cobs:







$\xi : [0, T] \to C$ — trajectory

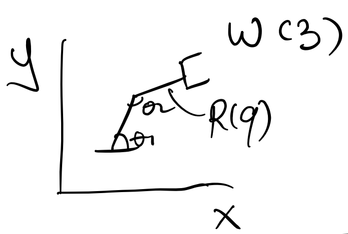
time, config. space

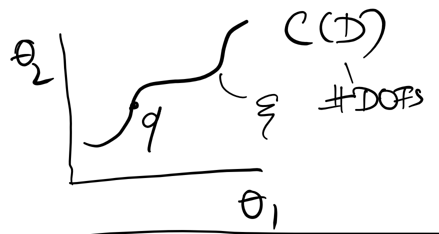
$\in \Xi$

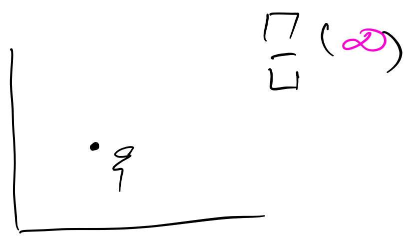
$\mathcal{U} : \Xi \to \mathbb{R}^+$ — cost functional

functions, scalars

$\mathcal{U}$: efficiency, smoothness, naturalness, legibility, obs avoidance!

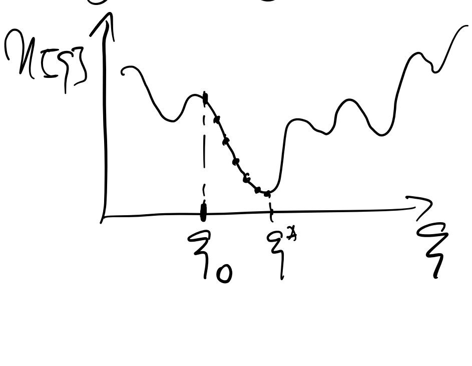


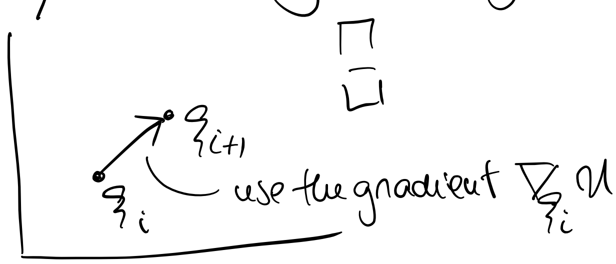




$$\boxed{\text{Find } \xi^* = \arg\min_{\xi \in \Xi} \mathcal{U}[\xi] \quad \text{s.t. } \xi(0) = q_s \quad \xi(T) = q_g}$$

Often: find ξ^* s.t. $\mathcal{U}[\xi^*] < \mu$ "good enough"





Calculus Background

$t \in \mathbb{R}, \quad t \in [0,T]; \quad q \in C \subset \mathbb{R}^D \quad D = \#DOFs$

$\xi : [0,T] \to C; \quad t \mapsto q; \quad t \mapsto \xi(t) = \begin{pmatrix} q^1(t) \\ q^2(t) \\ \vdots \\ q^D(t) \end{pmatrix}$ vector-valued function → vector calculus

if we discretize time

$\xi = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix}_{N \cdot D \times 1}$ $\quad \mathcal{U} : \square \to \mathbb{R}^+$

multivariate function
→ multivariate calculus

if we don't : $N = \infty$ → functional
→ calculus of variations

Derivative : $f(x) : \mathbb{R} \to \mathbb{R}$

$f'(x) = \lim_{\varepsilon \to 0} \dfrac{f(x+\varepsilon) - f(x)}{\varepsilon}$ (slope) (scalar)

example: $f(x) = 2x \quad f'(x) = \lim_{\varepsilon \to 0} \dfrac{2x + 2\varepsilon - 2x}{\varepsilon} = 2$

Partial derivative : assume other variables are constant

$f(x,y) : \mathbb{R}^2 \to \mathbb{R}$ (e.g. $\mathcal{U}$)

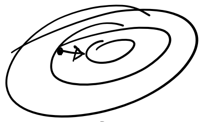
$\dfrac{\partial f}{\partial x}(x,y) = \lim_{\varepsilon \to 0} \dfrac{f(x+\varepsilon, y) - f(x,y)}{\varepsilon}$ (scalar)

example: $f(x,y) = 2x + y + xy$

$\dfrac{\partial f}{\partial x}(x,y) = \lim_{\varepsilon \to 0} \dfrac{\cancel{2x} + 2\varepsilon + \cancel{y} + \cancel{xy} + \varepsilon y - \cancel{2x} - \cancel{y} - \cancel{xy}}{\varepsilon}$

$= \lim_{\varepsilon \to 0} \dfrac{2\varepsilon + \varepsilon y}{\varepsilon} = 2 + y$

Gradient:



$f(x,y)$

$$\nabla_{(x,y)} f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{pmatrix}$$

(\nabla)

(vector) (direction of greatest rate of increase) (magnitude = slope in that dir.) (* see end)

example: $f(x,y) = 2x + y + xy \quad \nabla_{(x,y)} f(x,y) = \begin{pmatrix} 2+y \\ 1+x \end{pmatrix}$

Ordinary derivative of vector-valued function:

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix}$$ (e.g. ξ)

$$f'(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x} \\ \frac{\partial f_2}{\partial x} \end{pmatrix}$$ (vector)

example: $f(x) = \begin{pmatrix} x+2 \\ 2x \end{pmatrix} \quad f'(x) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

Vector by vector:

$$f(x,y) = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix}$$ (e.g. forward kinematics) $\phi(q) = \begin{pmatrix} x(q) \\ y(q) \\ z(q) \end{pmatrix}$

$$\frac{\partial f}{\partial (x,y)} = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix}$$ – Jacobian matrix

example: $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ y \end{pmatrix} \quad J = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$

$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad f(v) = Av \Rightarrow J = A$

$f(v) = v^T A v \Rightarrow f = v^T(A+A^T) \qquad \nabla = f^T$

(scalar valued multivariate

$\nabla = (A+A^T)v$

if $A = I$, $\nabla = 2v$

if A symmetric, $\nabla = 2Av$

$$(x,y)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (ax+cy \quad bx+dy)\begin{pmatrix} x \\ y \end{pmatrix} =$$

$$= ax^2 + cxy + bxy + dy^2$$

$$\nabla = \begin{pmatrix} 2a\,x + (b+c)y \\ (b+c)x + 2dy \end{pmatrix} = \underbrace{\begin{pmatrix} 2a & b+c \\ b+c & 2d \end{pmatrix}}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \underset{A}{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} + \underset{A^T}{\begin{pmatrix} a & c \\ b & d \end{pmatrix}}$$

2^{nd} order derivatives:

$f(x): \mathbb{R} \to \mathbb{R}$; $f'(x): \mathbb{R} \to \mathbb{R}$; $f''(x) = \frac{d f'(x)}{dx}$

$f(x,y): \mathbb{R}^2 \to \mathbb{R}$; $\nabla_{(x,y)} f : \mathbb{R}^2 \to \mathbb{R}^2$; $\nabla^2_{(x,y)} f$ = Jacobian of $\nabla_{(x,y)} f$

Hessian matrix

Functional gradient:

$f(g) \qquad g: \mathbb{R} \to \mathbb{R} \qquad f: (\mathbb{R} \to \mathbb{R}) \to \mathbb{R} \qquad$ (e.g. $\mathcal{U}$)

Note: different from $f(g(x)) = h(x) : \mathbb{R} \to \mathbb{R}$

↳ next lecture

Chain rule

$F(x) = f(g(x)) \quad F'(x) = f'(g(x)) \cdot g'(x)$

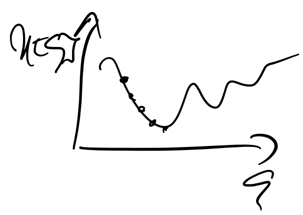
$y = g(x) \quad z = f(g(x)) \quad \frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$

multivariate: $F(x(\varepsilon), y(\varepsilon)) \quad \frac{dF}{d\varepsilon} = \frac{\partial F}{\partial x}\frac{dx}{d\varepsilon} + \frac{\partial F}{\partial y}\frac{dy}{d\varepsilon}$

Integration by parts

$$\int_a^b f g' \, dx = fg\Big|_a^b - \int_a^b g f' \, dx$$

Functional Gradient descent



$$\xi_{i+1} = \xi_i - \frac{1}{\alpha} \nabla_{\xi_i} U$$

Note: $\mathcal{H}$ - Hilbert space

↪ complete vector space + inner product

For now, assume $\underbrace{\langle \xi_1, \xi_2 \rangle = \int \xi_1(t)^T \xi_2(t) \, dt}_{\text{"Euclidean"}}$

- symmetry ✓
- linearity in 1st argument

$\langle a\xi_1, \xi_2 \rangle = a \langle \xi_1, \xi_2 \rangle$ ✓

$\langle \xi_1 + \xi_2, \xi_3 \rangle = \langle \xi_1, \xi_3 \rangle + \langle \xi_2, \xi_3 \rangle$ ✓

- pos. definiteness

$\langle g, g\rangle \geq 0$ ✓

$\langle g, g\rangle = 0 \Rightarrow g = 0$ ✓

$g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \end{bmatrix} \Rightarrow \langle g_1, g_2\rangle = g_1^T g_2$ "Euclidean"

<u>Euler-Lagrange</u>

If $\mathcal{U}[g] = \int_0^T F(t, g(t), g'(t))\, dt$

and $\langle g_1, g_2\rangle = \int g_1(t)^T g_2(t)\, dt$

then $\boxed{\nabla_g \mathcal{U}(t) = \frac{\partial F}{\partial g(t)}(t) - \frac{d}{dt}\frac{\partial F}{\partial g'(t)}(t)}$

Example:

$\mathcal{U}[g] = \frac{1}{2}\int \| g'(t)\|^2 dt$ $\left\{ \begin{array}{l} \text{shape? straight line} \\ \text{timing? constant velocity} \end{array} \right.$

$\nabla_g \mathcal{U}(t) = 0 - \frac{d}{dt} g'(t) = -g''(t)$

global min $\nabla_g \mathcal{U} = 0 \Rightarrow -g''(t) = 0 \Rightarrow g'(t) = a$

$\Rightarrow g(t) = at + b$

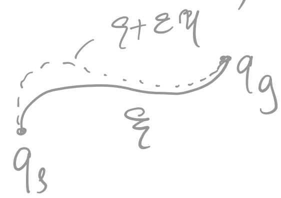
[determine a, b from $g(0), g(T)$]

Proof:

$f: \mathbb{R} \to \mathbb{R}$ $\quad f(x+\varepsilon) \approx f(x) + \varepsilon f'(x)$ $\quad$ 1st order Taylor

$\mathcal{U}: \mathcal{Q} \to \mathbb{R}^+$ $\quad \mathcal{U}[q+\varepsilon\eta] \approx \mathcal{U}[q] + \varepsilon \langle \nabla_q \mathcal{U}, \eta \rangle$

η: smooth disturbance



$\eta(0) = \eta(T) = 0$

$\varepsilon \in \mathbb{R}$

$$\Rightarrow \langle \nabla_q \mathcal{U}, \eta \rangle = \lim_{\varepsilon \to 0} \frac{\mathcal{U}[q+\varepsilon\eta] - \mathcal{U}[q]}{\varepsilon} \qquad (1)$$

def: $\langle \nabla_q \mathcal{U}, \eta \rangle = \int \nabla_q \mathcal{U}(t)^T \eta(t)\, dt \qquad (2)$

Let $\phi(\varepsilon) = \mathcal{U}[q+\varepsilon\eta]$

$$(1) \Rightarrow \langle \nabla_q \mathcal{U}, \eta \rangle = \lim_{\varepsilon \to 0} \frac{\phi(\varepsilon) - \phi(0)}{\varepsilon - 0} = \frac{d\phi}{d\varepsilon}(0)$$

$$= \frac{d\left(\int_0^T F(t, q(t)+\varepsilon\eta(t), q'(t)+\varepsilon\eta'(t))\, dt\right)}{d\varepsilon}(0)$$

exchange integration-differentiation (Leibnitz's Rule)

$$\int_0^T \frac{d}{d\varepsilon} F(t, \underbrace{q(t)+\varepsilon\eta(t)}_{x(\varepsilon)}, \underbrace{q'(t)+\varepsilon\eta'(t)}_{y(\varepsilon)})\, dt \quad (0)$$

chain rule (multivariate)

$$F(x(\varepsilon), y(\varepsilon)) \qquad \frac{dF}{d\varepsilon} = \frac{\partial F}{\partial x}^T \frac{dx}{d\varepsilon} + \frac{\partial F}{\partial y}^T \frac{dy}{d\varepsilon}$$

$$\int_0^T \frac{\partial F}{\partial x}^T \eta(t) + \frac{\partial F}{\partial y}^T \eta'(t)\, dt \quad (0)$$

$$\int_0^T \frac{\partial F}{\partial \varsigma(t)}(t, \varsigma(t), \varsigma'(t))^T \eta(t) + \frac{\partial F}{\partial \varsigma'(t)}(t, \varsigma(t), \varsigma'(t))^T \eta'(t)\, dt$$

compactly

$$\int_0^T \frac{\partial F}{\partial \varsigma}^T \eta + \frac{\partial F}{\partial \varsigma'}^T \eta' \, dt$$

integration by parts

$$\int_a^b f g' \, dx = f g \Big|_a^b - \int_a^b g f' \, dx$$

$$\int_0^T \frac{\partial F}{\partial \varsigma'}^T \eta' \, dt = \underbrace{\frac{\partial F}{\partial \varsigma'} \eta \Big|_0^T}_{\substack{0 \text{ because} \\ \eta(0)=\eta(T)=0}} - \int_0^T \frac{d}{dt} \frac{\partial F}{\partial \varsigma'} \eta \, dt$$

$$\Rightarrow \frac{d}{d\varepsilon} \phi(0) = \int_0^T \frac{\partial F}{\partial \varsigma}^T \eta - \frac{d}{dt} \frac{\partial F}{\partial \varsigma'} \eta \, dt$$

$$= \int_0^T \left[\frac{\partial F}{\partial \varsigma} - \frac{d}{dt} \frac{\partial F}{\partial \varsigma'} \right]^T \eta \, dt$$

$$(2): \quad = \int_0^T \nabla_\varsigma \mathcal{U}^T \eta \, dt \qquad \forall \eta$$

$$\Rightarrow \nabla_\varsigma \mathcal{U} = \frac{\partial F}{\partial \varsigma} - \frac{d}{dt} \frac{\partial F}{\partial \varsigma'}$$

Example 11: Covariant Hamiltonian Optimization for Motion Planning

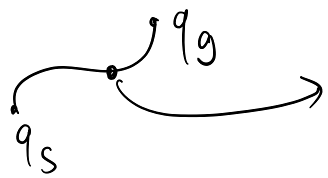
$U[\xi] = U_{obs}[\xi] + \lambda U_{smooth}[\xi]$

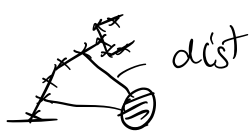
$c: W \to \mathbb{R}^+$ based on signed distance field

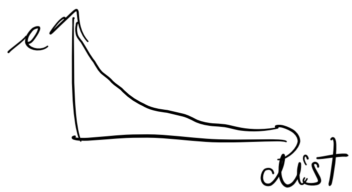
distance from closest obstacle in W

cheap, because in W

$$U_{obs}[\xi] = \int_t \int_b c(\phi_b(\xi(t))) \cdot \left\| \frac{d}{dt} \phi_b(\xi(t)) \right\| \, db\, dt$$







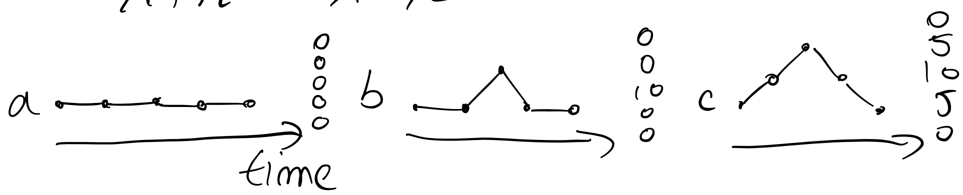
Note: can take gradient of U_{obs} without explicit representation of C_{obs} – uses Jacobian to pass info from workspace/world into C space

$$U_{smooth}[\xi] = \frac{1}{2} \int \| \xi'(t) \|^2 dt$$

Make it prob. complete via Hamiltonian Monte Carlo: sample some traj. momentum

Different Inner Products

$$\langle \xi_1, \xi_2 \rangle = \xi_1^T \xi_2$$



$$\|a-b\|^2 = (a-b)^T(a-b) = 100$$
$$\|a-c\|^2 = (a-c)^T(a-c) = 150$$
$$\Rightarrow \|a-b\|^2 < \|a-c\|^2 \quad \times$$

<u>Idea</u>: $\langle \xi_1, \xi_2 \rangle = \xi_1^T A \xi_2$ make c closer to a than b is

detour: $\mathcal{U}(\xi) = \sum_t \|q_{t+1} - q_t\|^2$

$$K = \begin{bmatrix} 1 & & & 0 \\ -1 & 1 & & \\ & \ddots & \ddots & \\ 0 & & -1 & 1 \\ & & & -1 \end{bmatrix} \quad K \cdot \begin{bmatrix} q_2 \\ \vdots \\ q_{N-1} \end{bmatrix} = \begin{bmatrix} q_2 \\ q_3 - q_2 \\ \vdots \\ q_{N-1} - q_{N-2} \\ -q_{N-1} \end{bmatrix} \quad e = \begin{bmatrix} -q_1 \\ 0 \\ \vdots \\ 0 \\ q_N \end{bmatrix}$$

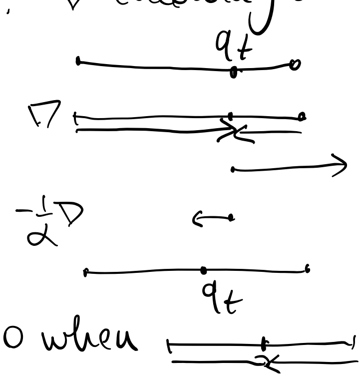
$$\mathcal{U}(\xi) = \frac{1}{2}(K\xi + e)^T(K\xi + e) = \frac{1}{2}\xi^T \underbrace{K^T K}_{A} \xi + \dots$$

$$\nabla_\xi \mathcal{U}(t) = -(q_{t+1} - q_t) + q_t - q_{t-1} = 2q_t - q_{t+1} - q_{t-1}$$

$$\nabla_\xi \mathcal{U} = A \cdot \xi \quad (+\text{ endpoint fixes})$$

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

note: ∇ encourages constant velocity:



$\langle \varsigma_1, \varsigma_2 \rangle = \varsigma_1^T A \varsigma_2$

$\|a-b\|^2 = (a-b)^T A (a-b) = (0,0,10,0,0) \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 10 \\ 0 \\ 0 \end{pmatrix}$

$(0,-10,20,-10,0) \begin{pmatrix} 0 \\ 0 \\ 10 \\ 0 \\ 0 \end{pmatrix} = \underline{200}$

$\|a-c\|^2 = (a-c)^T A (a-c) = (0,5,10,5,10) \begin{pmatrix} & \end{pmatrix} \begin{pmatrix} 0 \\ 5 \\ 10 \\ 5 \\ 0 \end{pmatrix}$

$= (-5, 10-10, -5+20-5, -10+10, -5) \begin{pmatrix} 0 \\ 5 \\ 10 \\ 5 \\ 0 \end{pmatrix}$

$= (-5, 0, 10, 0, -5) \begin{pmatrix} 0 \\ 5 \\ 10 \\ 5 \\ 0 \end{pmatrix} = 0+0+100+0+0 = \underline{\underline{100}}$

$\Rightarrow \|a-c\|^2 < \|a-b\|^2$

Impact on gradient:

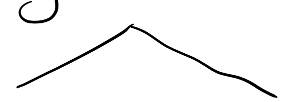
$\mathcal{U}[\varsigma] \approx \mathcal{U}[\varsigma_i] + \langle \nabla_\varsigma^I \mathcal{U}, \varsigma - \varsigma_i \rangle_I$

$\mathcal{U}[\varsigma] \approx \mathcal{U}[\varsigma_i] + \langle \nabla_\varsigma^A \mathcal{U}, \varsigma - \varsigma_i \rangle_A$

$\Rightarrow \nabla_\varsigma^I \mathcal{U}^T I (\varsigma - \varsigma_i) = \nabla_\varsigma^A \mathcal{U}^T A (\varsigma - \varsigma_i), \forall \varsigma$

$\Rightarrow \nabla_\varsigma^I \mathcal{U}^T = \nabla_\varsigma^A \mathcal{U}^T A \Rightarrow \boxed{\nabla_\varsigma^A \mathcal{U} = A^{-1} \nabla_\varsigma^I \mathcal{U}}$

propagate Euclidean gradient through A^{-1}

$\nabla_\varsigma^I \mathcal{U} =$ $\nabla_\varsigma^A \mathcal{U} =$ 

turns trajectory into elastic band

Relation to Newton Method / Alternative derivation:

$$\xi_{i+1} = \underset{\xi}{\operatorname{argmin}} \underbrace{\mathcal{U}[\xi_i] + \nabla_\xi^{I} \mathcal{U}^T (\xi - \xi_i)}_{\text{first order TSE}} + \underbrace{\frac{\alpha}{2} \|\xi - \xi_i\|_A^2}_{\text{stay close to } \xi_i}$$

$\nabla = 0$: $\quad 0 + \nabla_\xi^{I} \mathcal{U} + \alpha \cdot A(\xi - \xi_i) = 0$

$$\xi - \xi_i = -\frac{1}{\alpha} A^{-1} \nabla_\xi^{I} \mathcal{U}$$

$$\boxed{\xi = \xi_i - \frac{1}{\alpha} A^{-1} \nabla_\xi^{I} \mathcal{U}}$$

Newton's Method: replace A with Hessian

(A is Hessian of $\mathcal{U}_{smooth}$ only)

(*) Gradient direction = steepest direction

$$\max_{\Delta : \Delta x^2 + \Delta y^2 \leq \varepsilon} f((x,y)+\Delta)$$

$$f((x,y)+\Delta) \approx \underbrace{f(x,y)}_{\text{does not depend on } \Delta} + \underbrace{\frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y}_{\Delta^T \nabla}$$

$$\max_{\Delta : \|\Delta\| \leq \sqrt{\varepsilon}} \Delta^T \nabla \rightarrow \text{align } \Delta \text{ with } \Delta$$

$$\Delta = \frac{\nabla}{\|\nabla\|} \cdot \sqrt{\varepsilon}$$