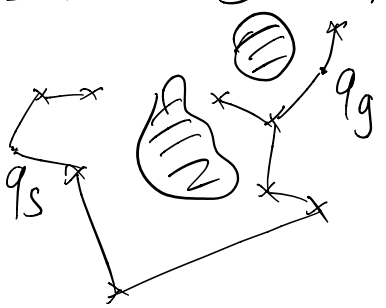
Trajectory Optimization

Last time: Bi-RRT



⊕ works in high dim spaces
⊕ fast (ms-s.)
⊖ not complete
⊖ not optimal → shortcutting
still not optimal
no longer as fast

"Optimal motion planning": randomized sampling + asyptotic optimality

Eg RRT* [Karaman & Frazzoli, 2010]
- pick parent by considering cost-to-come
- rewire: $\forall q \in Near(q_{sample})$
  if less costly to reach q from q_{sample}
  ↳ switch q's parent to q_{sample}

⊕ asymptotically optimal
⊖ slow: optimize to everywhere - wasted effort

Today: prob. complete, fast, (locally) optimal
Idea: start with simple path, let obstacles push you out of the way

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

Problem Statement

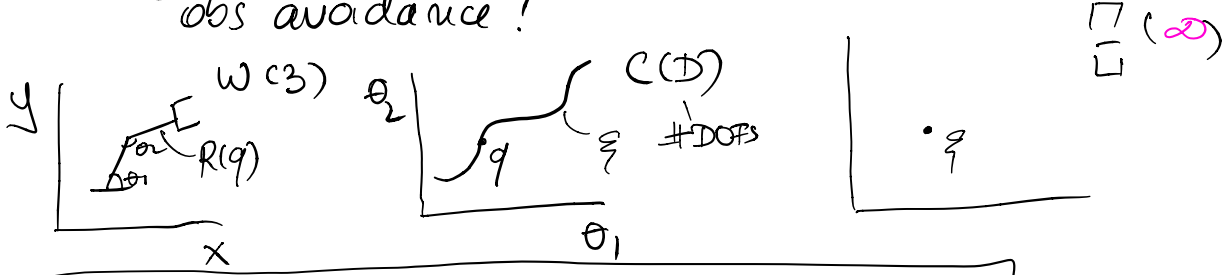
$\xi : [0, T] \to \mathcal{C}$ — trajectory $\in \Pi$

time; config. space

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

functions; scalars

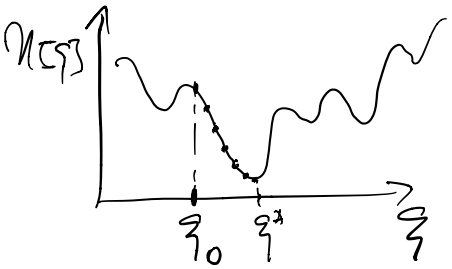
$\mathcal{U}$: efficiency, smoothness, naturalness, legibility, obs avoidance!

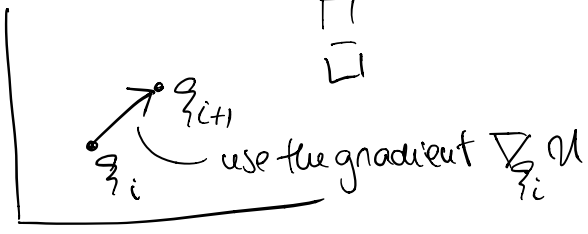


Find $\xi^* = \underset{\xi \in \Pi}{\operatorname{argmin}} \mathcal{U}[\xi]$

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

$\to$ vector calculus

if we discretize time

$$\xi = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix}_{N \cdot D}$$

1

$\mathcal{U} : \square \to \mathbb{R}^+$

multivariate function

$\to$ multivariate calculus

if we don't : $N = \infty$ $\to$ functional

$\to$ calculus of variations

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

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

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

Partial derivative : assume other variables are constant

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

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

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

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

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

Gradient:

$f(x,y)$

$\nabla_{(x,y)} f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{pmatrix}$ (vector) (direction of greatest rate of increase) (magnitude = slope in that dir.) (* see end)

∇ = nabla

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

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

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

(scalar valued multivariate)

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

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

if A symmetric, $\nabla = 2Av$

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

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

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

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

2nd order derivatives:

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

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

Hessian matrix

Functional gradient:

$J(g)$ $\quad g: \mathbb{R} \to \mathbb{R}$ $\quad J: (\mathbb{R} \to \mathbb{R}) \to \mathbb{R}$ $\quad$ (e.g. $\mathcal{U}$)

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

↳ next lecture

Chain rule

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

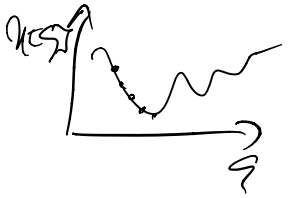
$y = g(x) \quad z = f(g(x)) \quad \frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$

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



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

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

↳ complete vector space + inner product

for now, assume $\langle \xi_1, \xi_2 \rangle = \int \xi_1(t)^T \xi_2(t) \, dt$

"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

$\langle \xi, \xi \rangle \geq 0$ ✓

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

$\xi = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \end{bmatrix} \Rightarrow \langle \xi_1, \xi_2 \rangle = \xi_1^T \xi_2$ "Euclidean"

Euler-Lagrange

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

and $\langle \xi_1, \xi_2 \rangle = \int \xi_1(t)^T \xi_2(t)\, dt$

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

Example:

$\mathcal{U}[\xi] = \frac{1}{2} \int \|\xi'(t)\|^2 dt$ { shape? straight line; timing? constant velocity

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

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

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

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

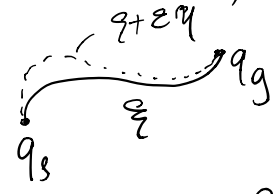
Proof:

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

$\mathcal{U}: \frac{\Gamma}{L} \to \mathbb{R}^+$ $\mathcal{U}[\xi + \varepsilon\eta] \approx \mathcal{U}[\xi] + \varepsilon \langle \nabla_\xi \mathcal{U}, \eta \rangle$

η: smooth disturbance $\quad \eta(0) = \eta(T) = 0$

$\varepsilon \in \mathbb{R}$



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

dy: $$\langle \nabla_\xi \mathcal{U}, \eta \rangle = \int \nabla_\xi \mathcal{U}(t)^T \eta(t)\, dt \quad (2)$$

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

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

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

exchange integration - differentiation (Leibnitz's Rule)

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

chain rule (multivariate)

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

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

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

compactly

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

Integration by parts

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

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

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

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

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

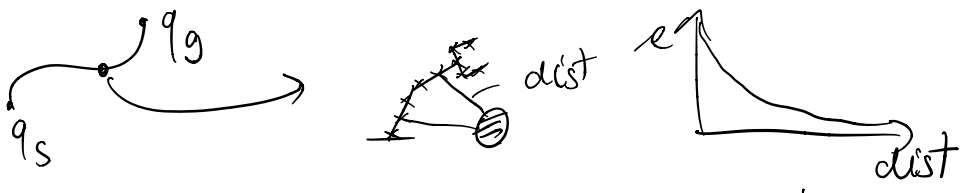
$$\Rightarrow \nabla_\varsigma \mathcal{U} = \frac{\partial F}{\partial \varsigma} - \frac{d}{dt}\frac{\partial F}{\partial \varsigma'}$$

Example 11: Covariant Hamiltonian Optimization for Motion Planning

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

$c: \mathcal{W} \to \mathbb{R}^+$ based on signed distance field
distance from closest obstacle in $\mathcal{W}$
clearp, because in $\mathcal{W}$

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



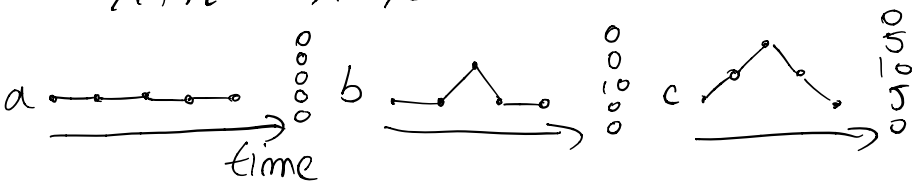
Note: can take gradient of U_{obs} without explicit representation of $c(\phi_b)$ – uses Jacobian to pass info from workspace/world into cspace

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

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

Different Inner Products

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



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

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

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

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

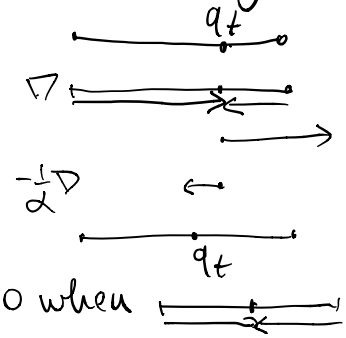
$U(q) = \frac{1}{2}(Kq + e)^T(Kq + e) = \frac{1}{2} q^T \underbrace{K^T K}_{A} q + \dots$

$\nabla_q U(t) = -(q_{t+1} - q_t) + q_t - q_{t-1} = 2q_t - q_{t+1} - q_{t-1}$

$\nabla_q U = A \cdot q$ (+ endpoint fixes)

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

note: ∇ encourages constant velocity:



$\langle s_1, s_2 \rangle = s_1^T A s_2$

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

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

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

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

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

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

Impact on gradient:

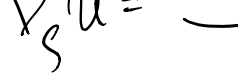
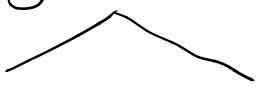
$U[s] \approx U[s_i] + \langle \nabla_s^I U, s - s_i \rangle_I$

$U[s] \approx U[s_i] + \langle \nabla_s^A U, s - s_i \rangle_A$

$\Rightarrow \nabla_s^I U^T I (s - s_i) = \nabla_s^A U^T A (s - s_i), \forall s$

$\Rightarrow \nabla_s^I U^T = \nabla_s^A U^T A \Rightarrow \boxed{\nabla_s^A U = A^{-1} \nabla_s^I U}$

propagate Euclidean gradient through A^{-1}

$\nabla_s^I U =$  $\nabla_s^A U =$ 

turns trajectory into elastic band

Relation to Newton Method / Alternative derivation:

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

$$\nabla = 0: \quad 0 + \nabla_\xi^T U + \alpha \cdot A(\xi - \xi_i) = 0$$

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

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

Newton's Method: replace A with Hessian

(A is Hessian of U smooth only)

(*) Gradient direction = steepest direction

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

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

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

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