Last Time

- Newton-Euler Dynamics
- SE(3)
- Quadrotors
- Airplanes

Today:

- HW 1 Q+A
- Optization Review

HW 1 Notes:

* Exponential:

$$\exp : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n} \;/\; \mathbb{H} \to \mathbb{H}$$

- We also commonly define:

$$\mathrm{Exp} : \mathbb{R}^3 \to SO(3) \;/\; \mathbb{R}^3 \to \mathbb{H}$$

(basically composing hat map w/ exp)

* Lie Midpoint

- For 2^{nd} order, we need $Q(t+h)$ to match Taylor expansion to 2^{nd} order:

$$Q(t+h) \approx Q(t) + h\dot{Q}(t) + \frac{1}{2}h^2\ddot{Q}(t)$$

$$\approx Q(t) + hQ_n\hat{\omega}_n + \frac{1}{2}h^2\left[Q_n\hat{\omega}_n^2 + Q\hat{\dot{\omega}}_n\right]$$

$$(+\ \text{H.O.T.})$$

- Now, look at RK-MK step:

$$Q_{n+1} = Q_n \exp(h\hat{\omega}_m)$$
$$= Q_n \exp\left(h\left(\hat{\omega}_n + \tfrac{1}{2}h\hat{\dot{\omega}}_n\right)\right)$$

- Now Taylor expand exp:

$$\exp(x) \approx I + x + \tfrac{1}{2}x^2 + \ldots$$

$$\Rightarrow Q_{n+1} \approx Q_n\left[I + h\left(\hat{\omega}_n + \tfrac{1}{2}h\hat{\dot{\omega}}_n\right) + \tfrac{1}{2}h^2\hat{\omega}_n^2 + H.O.T.\right]$$
$$\approx Q_n + hQ_n\hat{\omega}_n + \tfrac{1}{2}h^2 Q_n\left[\hat{\omega}_n^2 + \hat{\dot{\omega}}_n\right]$$

Optimization Review Pt. 1:

* Notation:

- Given $f(x): \mathbb{R}^n \to \mathbb{R}$

$$\underbrace{\frac{\partial f}{\partial x}}_{\text{"gradient"}} \text{ is a } \underline{\text{row}} \text{ vector}$$

- This is because $\frac{\partial f}{\partial x}$ maps $\Delta x \to \Delta f$:

$$f(x+\Delta x) \approx f(x) + \underbrace{\frac{\partial f}{\partial x}}_{1\times n}\underbrace{\Delta x}_{n\times 1}$$

- Similarly, given $g(y): \mathbb{R}^m \to \mathbb{R}^n$

$$\underbrace{\frac{\partial g}{\partial y}}_{\text{"Jacobian"}} \in \mathbb{R}^{n\times m} \quad \text{because} \quad g(y+\Delta y) \approx g(y) + \frac{\partial g}{\partial y}\Delta y$$

- These conventions are important because they make the chain rule work!

$$f(g(y+\Delta y)) \approx f(g(y)) + \left.\frac{\partial f}{\partial g}\right|_{g(y)} \left.\frac{\partial g}{\partial y}\right|_{y} \Delta y$$

- For convenience, we also define:

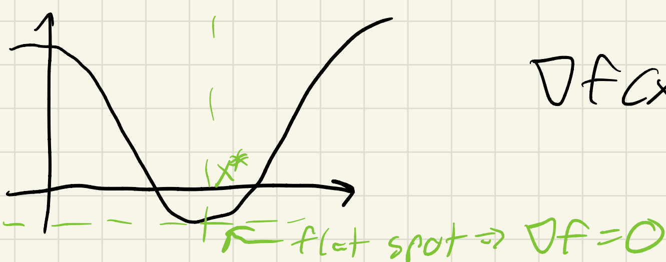
$$\nabla f(x) = \left(\frac{\partial f}{\partial x}\right)^T \in \mathbb{R}^{n\times 1} \quad \underline{\text{column}} \text{ vector}$$

$$\underbrace{\nabla^2 f(x) = \frac{\partial^2 f}{\partial x^2}}_{\text{"Hessian"}} = \frac{\partial}{\partial x}\left(\nabla f(x)\right) \in \mathbb{R}^{n\times n}$$

* Minimization

$$\min_x f(x), \quad f(x): \mathbb{R}^n \to \mathbb{R}$$

- First-order necessary conditions:



$$\nabla f(x^*) = 0$$

- Flat spot could be a min, max, or a saddle.

- Second-Order Sufficient Conditions:

$$\nabla f(x^*) = 0 \quad , \quad \nabla^2 f(x) > 0$$

Positive definite

Positive curvature (function is increasing locally in all directions)

"weakly"

- Convex functions have $\nabla^2 f(x) \geq 0 \ \forall x$

- Strongly convex functions have $\nabla^2 f(x) > 0$

  - Strongly convex functions have a single unique global minimizer x^*

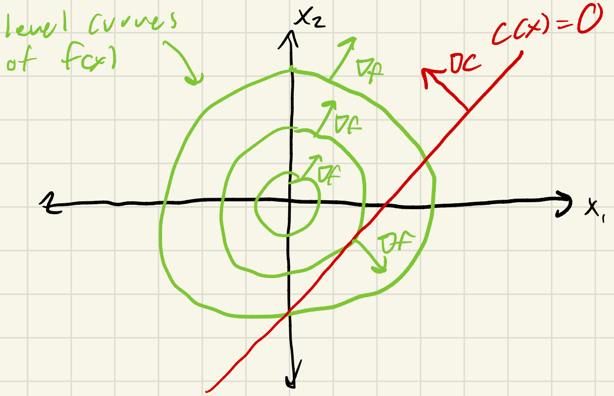
- Weakly convex functions have a global minimum value of f but can have set-valued minima (e.g. valley with a flat bottom).

* We will assume functions are at least C^2 smooth (2^{nd} derivatives are continuous). Not strictly necessary, but makes math easier.

* Minimization with Equality Constraints

$$\min_x f(x) \quad , \quad f(x) : \mathbb{R}^n \to \mathbb{R}$$
$$\text{s.t. } c(x) = 0 \quad , \quad c(x) : \mathbb{R}^n \to \mathbb{R}^m$$

- First-order necessary conditions:
  1) Need $\nabla f(x) = 0$ in <u>free</u> directions
  2) Need $c(x) = 0$



∇c = normal vector to $c(x) = 0$

- Any non-zero component of ∇f must be normal to the constraint surface/manifold

$$\Rightarrow \quad \nabla f + \underbrace{\lambda}_{\text{scalar "Lagrange Multiplier" / "Dual Variable"}} \nabla c = 0 \quad , \quad \lambda \in \mathbb{R}$$

- In general:

$$\frac{\partial f}{\partial x} + \lambda^T \frac{\partial c}{\partial x} = 0$$

- Based on this condition, we define the "Lagrangian"

$$L(x,\lambda) = f(x) + \lambda^T c(x)$$

Such that:

$$\left.\begin{aligned} \nabla_x L(x,\lambda) &= \nabla f + \left(\frac{\partial f}{\partial c}\right)^T \lambda = 0 \\ \nabla_\lambda L(x,\lambda) &= c(x) = 0 \end{aligned}\right\} \text{"KKT conditions"}$$

- Second-Order Sufficiency Conditions:

$$\nabla^2 f(x) > 0 \text{ in } \underline{\text{free}} \text{ directions}$$

* Solving Optimization Problems
  - KKT conditions define a root-finding problem:

$$\nabla L(x,\lambda) = 0$$

  - Classic solution technique: Newton's Method
    1) Start with a guess for x^*, λ^*
    2) Linearize ∇L about guess:

$$\nabla_x L(x+\Delta x, \lambda+\Delta\lambda) \approx \nabla_x L(x,\lambda) + \frac{\partial^2 L}{\partial x^2}\Delta x + \underbrace{\frac{\partial^2 L}{\partial x \partial \lambda}}_{\left(\frac{\partial c}{\partial x}\right)^T}\Delta\lambda$$

$$\nabla_\lambda L(x+\Delta x, \lambda+\Delta\lambda) \approx c(x) + \frac{\partial c}{\partial x}\Delta x$$

3) Set linearization $= 0$ and solve for $\Delta x, \Delta\lambda$

$$\underbrace{\begin{bmatrix} \nabla^2 L & \frac{\partial c}{\partial x}^T \\ \frac{\partial c}{\partial x} & 0 \end{bmatrix}}_{\text{"KKT System"}} \begin{bmatrix} \Delta x \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} -\nabla_x L \\ -c(x) \end{bmatrix}$$

4) Update $x \leftarrow x + \Delta x$, $\lambda \leftarrow \lambda + \Delta\lambda$ and go to 2) until convergence.

- Some additional tricks are typically added to ensure convergence on difficult problems, but we won't need them.

* Gauss-Newton Method:

$$\frac{\partial^2 L}{\partial x} = \frac{\partial^2 f}{\partial f} + \underbrace{\frac{\partial}{\partial x}\left[\left(\frac{\partial c}{\partial x}\right)^T \lambda\right]}_{\text{3rd-rank tensor}} \leftarrow \text{"constraint curvature"}$$

- 3rd-rank tensor is expensive/annoying to compute

- We often drop 2nd term. Called "Gauss-Newton".

- Still does almost as well as full Newton