$$\boxed{\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x)=0 \end{array}}$$

$g: \mathbb{R}^n \to \mathbb{R}^m$

CONSTRAINTS:

$g(x)=0 \qquad g: \mathbb{R}^n \to \mathbb{R}$

$$\mathcal{L}(x,\lambda) = f(x) + \lambda^T g(x)$$

FONC: $\boxed{\begin{array}{l} \nabla_x \mathcal{L}(x^*,\lambda^*)=0 \\ g(x^*) = 0 \end{array}}$

[MORE GENERAL: KKT-CONDITIONS]

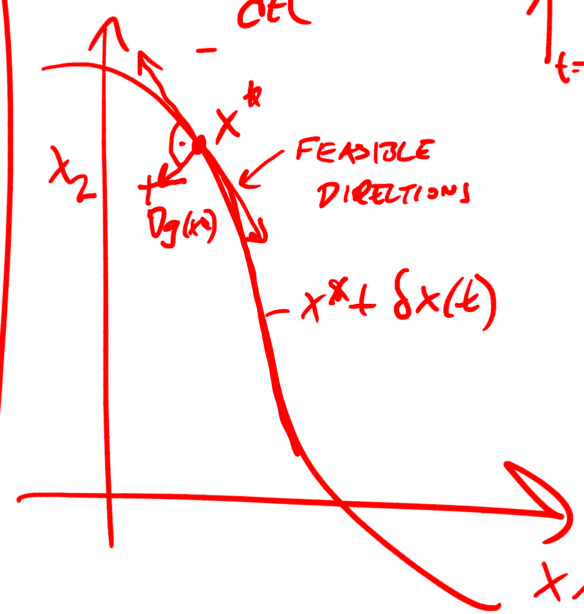
$g(x^*)=0$

$g(x^*+\delta x(t))=0$

$$\frac{d}{dt}\Big(\qquad \Big)\Big|_{t=0} = \nabla g(x^*)^T \underbrace{\frac{d\,\delta x(t)}{dt}}_{Z \cdot v} \overset{!}{=} 0$$

LICQ $\Leftrightarrow$ ∇g FULL RANK

$\nabla g^T \cdot Z = 0$ Z BASIS OF NULLSPACE/KERNEL OF ∇g^T



SONC/SOSC: $Z^T \nabla^2 \mathcal{L}(x^*,\lambda^*) Z \succeq 0$

$\underline{\succ 0}$

$Z \in \mathbb{R}^{n \times (n-m)}$

CH 2, END: STABILITY UNDER PERTURBATIONS

THM 8: SOSC (+ LICQ +STRICT COMPLEMENTARITY) IMPLIES STABILITY, I.E. SOLUTION MAP IS DIFFERENTIABLE AT $\delta=0$

$$\delta=(\delta_f,\delta_g,\delta_h) \longrightarrow \boxed{\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x)+\delta_f^T\cdot x \\ \text{s.t.} \quad & g(x)+\delta_g=0 \\ & h(x)+\delta_h\leq 0 \end{aligned}} \longrightarrow \begin{matrix} x^*(\delta) \\ \lambda^*(\delta) \\ \mu^*(\delta) \end{matrix}$$

"SOLUTION MAP"

$$\frac{d}{d\delta}\begin{bmatrix} x^*(\delta) \\ \tilde{\lambda}^*(\delta) \end{bmatrix}\Bigg|_{\delta=0} = -\underbrace{\left(\begin{array}{c|c} \nabla_x^2\mathcal{L} & \nabla\tilde{g} \\ \hline \nabla\tilde{g}^T & 0 \end{array}\right)^{-1}}_{\text{KKT-MATRIX}}\Bigg|_{\delta=0}$$

$$\tilde{\lambda}=\begin{bmatrix} \lambda \\ \mu_i \end{bmatrix}_{i\in\mathcal{A}}$$

$$\tilde{g}=\begin{bmatrix} g \\ h_i \end{bmatrix} \quad "$$

$$\delta=\begin{bmatrix} \delta_f \\ \tilde{\delta} \end{bmatrix} \qquad \tilde{\delta}=\begin{bmatrix} \delta_g \\ \delta_{h_i} \end{bmatrix} \quad "$$

$$\boxed{\left(\begin{array}{c|c} \nabla^2\mathcal{L} & \\ \hline & 0 \end{array}\right) \text{ INV}} \Leftarrow \left[\begin{matrix}\text{LICQ \&} \\ \text{SOSC}\end{matrix}\right]$$

$$\Rightarrow \text{LICQ}$$

CH 3 : NEWTON-TYPE OPTIMIZATION ALGORITHMS

3.1.

$$\boxed{\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x)=0 \end{array}} \quad g: \mathbb{R}^n \to \mathbb{R}^m \quad \boxed{\begin{array}{l} \nabla_x \mathcal{L}(x,\lambda) = 0 \\ g(x) = 0 \end{array}} \text{ KKT}$$

$$w = \begin{bmatrix} x \\ \lambda \end{bmatrix} \begin{matrix} n \\ m \end{matrix} \qquad F(w) = \begin{bmatrix} \nabla_x \mathcal{L} \\ g \end{bmatrix} \begin{matrix} n \\ m \end{matrix} \qquad F: \mathbb{R}^{(n+m)} \to \mathbb{R}^{(n+m)}$$

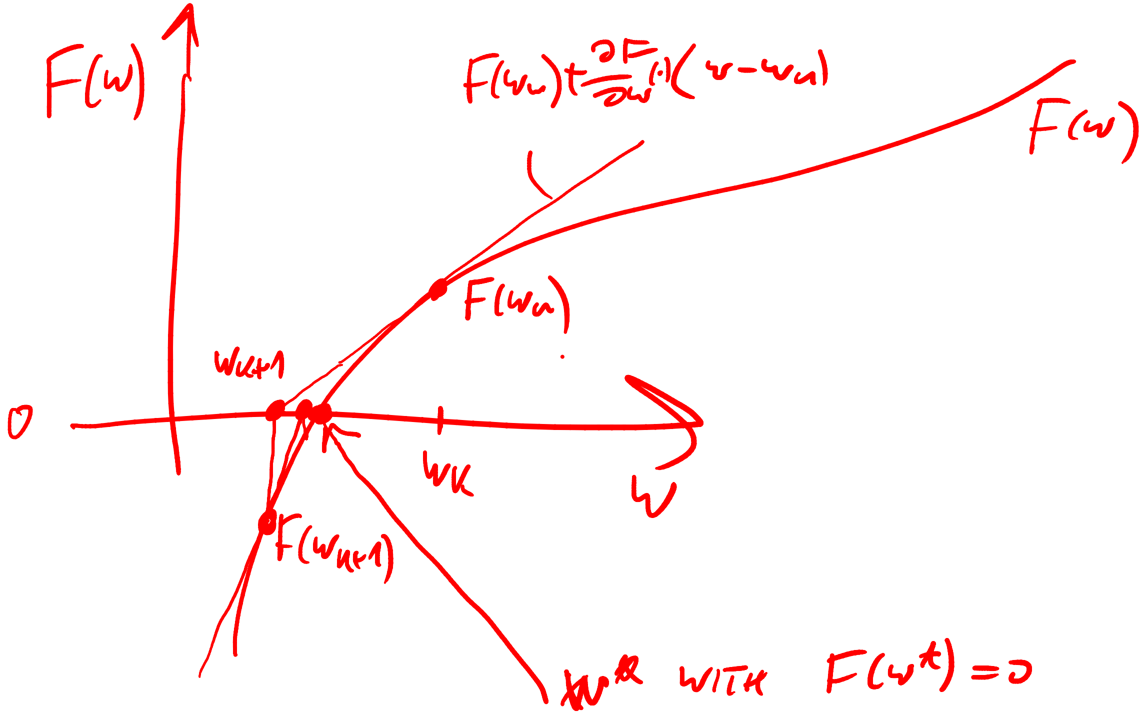
NEWTON'S METHOD: FIND w^* WITH $F(w^*)=0$

BY SUCCESSIVE LINEARIZATION AT "LINEARIZATION POINTS"

w_k, $w_k \to w^*$.

$$\boxed{F(w_k) + \frac{\partial F}{\partial w}(w_k)\left(\underline{\underline{w}} - w_k\right) = 0} \qquad \frac{\partial F}{\partial w} \in \mathbb{R}^{(n+m)\times(n+m)}$$

$$w_{k+1} = w_k - \left(\frac{\partial F}{\partial w}(w_k)\right)^{-1} \cdot F(w_k)$$



$$F(w_k) + \frac{\partial F}{\partial w}(\cdot)(v - w_k)$$

$$\begin{bmatrix}\nabla_x \mathcal{L}\\ g(x)\end{bmatrix} + \begin{bmatrix}\nabla_x^2\mathcal{L} & \nabla g\\ \nabla g^T & 0\end{bmatrix}\begin{bmatrix}x - x_k\\ \lambda - \lambda_k\end{bmatrix}$$

$$\nabla_x \mathcal{L} = \nabla f + \nabla g \cdot \lambda_k$$

$$\begin{bmatrix}\nabla f(x_k)\\ g(x_k)\end{bmatrix} + \underbrace{\begin{bmatrix}\boxed{\nabla_x^2\mathcal{L}} & \nabla g\\ \nabla g^T & 0\end{bmatrix}}_{= \text{KKT-MATRIX}}\begin{bmatrix}x - x_k\\ \lambda\end{bmatrix} \overset{!}{=} 0 \quad (*)$$

$$B_k \approx \nabla_x^2 \mathcal{L}$$

ALGO 1:

- GUESS x_0, λ_0
- $k = 0$
- WHILE $\left\|\begin{matrix}\nabla\mathcal{L}\\ g\end{matrix}\right\| \geq \varepsilon$ DO
  - GET B_k
  - SOLVE $(*)$ TO GET x_{k+1}, λ_{k+1}
  - $k = k+1$

3.1.1 QUADRATIC MODEL INTERPRETATION

ORIGINAL NLP

$$\boxed{\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \end{aligned}}$$

QP AT (x_k, λ_k):

$$\boxed{\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \nabla f(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T B_k (x - x_k) \\ \text{s.t.} \quad & g(x_k) + \nabla g(x_k)^T \cdot (x - x_k) = 0 \end{aligned}}$$

$B_k \approx \nabla^2 \mathcal{L}$

KKT-CONDITIONS OF QP:

$$\mathcal{L}^{QP}(x, \lambda) = f^{QP}(x) + g^{QP}(x)^T \cdot \lambda$$

KKT OF QP EQUIVALENT TO NEWTON-STEP EQUATION

$$0 = \nabla_x \mathcal{L}^{QP} = \nabla f(x_k) + B_k \cdot (x - x_k) + \nabla g(x_k) \cdot \lambda$$

$$0 = g^{QP} = g(x_k) + \nabla g(x_k)^T (x - x_k)$$

$$\boxed{\begin{bmatrix} \nabla f \\ g \end{bmatrix} + \begin{bmatrix} \boxed{B_k} & \nabla g \\ \nabla g^T & 0 \end{bmatrix} \begin{bmatrix} x - x_k \\ \lambda \end{bmatrix} = 0}$$

NEWTON-TYPE METHOD

SEQUENTIAL-QUADRATIC-PROGRAMMING (IMPORTANT FOR INEQUALITIES)

$$\underline{\underline{B_k = ?}}$$

RELATION TO DISCRETE TIME OPTIMAL CONTROL

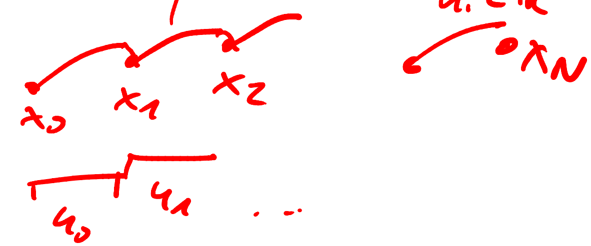
$$\boxed{\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x)=0 \end{array}} \iff \boxed{\begin{array}{ll} \min\limits_{y\in\mathbb{R}^n} & \phi(y) \\ \text{s.t.} & G(y)=0 \end{array}}$$

NLP

$$\boxed{\begin{array}{ll} \underset{y}{\text{minimize}} & \sum_{i=0}^{N-1} L(x_i,u_i) + E(x_N) \\ \text{s.t.} & f(x_i,u_i) - x_{i+1} = 0, \quad i=0,\dots,N-1 \\ & \bar{x}_0 - x_0 = 0 \end{array}}$$

OCP

$x_{i+1} = f(x_i,u_i)$ $\qquad x_i \in \mathbb{R}^{n_x}$, $u_i \in \mathbb{R}^{n_u}$



$$\phi(y) \overset{?}{=} \sum_{i=0}^{N-1} L(\cdot,\cdot) + E(x_N)$$

$$G(y) = \begin{bmatrix} f(x_0,u_0) - x_1 \\ f(x_1,u_1) - x_2 \\ \vdots \\ f(x_{N-1},u_{N-1}) - x_N \\ \bar{x}_0 - x_0 \end{bmatrix}$$

$$\boxed{\begin{array}{c} \nabla_y G(y)^T = \dots \\ \downarrow \\ \text{LQR} \dots \end{array}}$$

$$y = \begin{bmatrix} x_0 \\ u_0 \\ x_1 \\ u_1 \\ \vdots \\ x_{N-1} \\ u_{N-1} \\ x_N \end{bmatrix} \in \mathbb{R}^{N(n_x+n_u)+n_x}$$

3.1.2 EXACT NEWTON'S METHODS

$$B_k = \nabla_x^2 \mathcal{L}(x_k, \lambda_k)$$

3.1.3 GENERALIZED GAUSS-NEWTON-METHOD (GN)

ONLY FOR $\boxed{f(x) = \frac{1}{2}\|R(x)\|_2^2} = \frac{1}{2}\sum_{i=1}^{n_R}(R_i(x))^2$

$R: \mathbb{R}^n \to \mathbb{R}^{n_R}$

$$B_k = \nabla R(x_k) \cdot \nabla R(x_k)^T \in \mathbb{R}^{n \times n}$$

$\frac{\partial R}{\partial x} = \nabla R^T \in \mathbb{R}^{n_R \times n}$

$$\nabla_x^2 \mathcal{L} = \nabla^2 f + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x)$$

$$= \underbrace{\nabla R \nabla R^T}_{= B_k} + \sum_{j=1}^{n_R} R_j(x) \cdot \nabla_x^2 R_j(x)$$

$\approx B_k$ (IF R_j & λ_i ARE SMALL)

$\Uparrow$ IF $\|R(x_k)\|$ IS SMALL

$B_k \succeq 0$

$\nabla f = \sum_{i=1}^{n_R} R_i(x) \cdot \nabla R_i(x) = \nabla R \cdot R$

$\nabla^2 f =$

G-N IS GOOD IF "FIT" IS GOOD (RESIDUAL IS SMALL)

QP - INTERPRETATION FOR GN-METHOD

$$\boxed{\begin{array}{ll} \min & \nabla f(x_k)^T (x-x_k) + \frac{1}{2}(x-x_k)^T B_k (x-x_k) \\ \text{s.t.} & \dots \end{array}} \iff \boxed{\begin{array}{ll} \min\limits_x & \frac{1}{2}\|R(x_k) + \nabla R(x_k)^T (x-x_k)\|_2^2 \\ \text{s.t.} & g(x_k) + \nabla g(x_k)^T (x-x_k) = 0 \end{array}}$$

QP IS CONVEX !

$$\boxed{\begin{array}{l} \nabla f = \nabla R(x_k) \cdot R(x_k) \\ B_k = \nabla R \; \nabla R^T \end{array}}$$

NEWTON'S METHOD: QUADRATIC RATE OF CONVERGENCE

$$\|w_{k+1} - w^*\| \leq \frac{\omega}{2} \|w_k - w^*\|^2$$

G-N: LINEAR CONVERGENCE

$$\|w_{k+1} - w^*\| \leq K \|w_k - w^*\|$$

$0 \leq K < 1$

GOOD ENOUGH IF $K \ll 1$

0.1

0.01

0.001

0.00001

0.0000000001