

NONLINEAR PROGRAMMING 2

NEWTON-TYPE OPTIMIZATION

①

MOTIVATION: NLP ARISING FROM OPTIMAL CONTROL PROBLEMS (OCP-NLP)

CONTROL TRAJECTORY $(u_0, u_1, \dots, u_{N-1}) \in \mathbb{R}^{N \cdot m_u}$

STATE TRAJECTORY $(x_0, x_1, \dots, x_{N-1}, x_N) \in \mathbb{R}^{(N+1) \cdot n_x}$

NLP VARIABLES $w = (x_0, u_0, \dots, u_{N-1}, x_N) \in \mathbb{R}^{N(m_u + n_x) + n_x}$

$$\begin{aligned}
 & \min_w \sum_{k=0}^{N-1} L_k(x_k, u_k) + E(x_N) \quad (=: F(w)) \\
 & \text{s.t.} \quad 0 = \bar{x}_0 - x_0 \\
 & \quad \quad 0 = f_k(x_k, u_k) - x_{k+1} \quad (\text{SYSTEM DYNAMICS}) \\
 & \quad \quad 0 \geq h_k(x_k, u_k) \quad k=0, \dots, N-1 \\
 & \quad \quad 0 \geq r(x_N)
 \end{aligned}$$

(OCP-NLP)

SUMMARIZE AS

$$\begin{aligned}
 & \min_{w \in \mathbb{R}^n} F(w) \\
 & \text{s.t.} \quad G(w) = 0 \\
 & \quad \quad H(w) \leq 0
 \end{aligned}$$

(NLP)

$$\text{WITH } G(w) = \begin{bmatrix} \bar{x}_0 - x_0 \\ f_0(x_0, u_0) - x_1 \\ \vdots \\ f_{N-1}(x_{N-1}, u_{N-1}) - x_N \end{bmatrix} \in \mathbb{R}^m$$

$$\text{AND } H(w) = \begin{bmatrix} h_0(x_0, u_0) \\ \vdots \\ h_{N-1}(x_{N-1}, u_{N-1}) \\ r(x_N) \end{bmatrix}$$

(2)

KKT - OPTIMALITY CONDITIONS

IF w^* LOC. MIN. WITH (LICQ) THEN EXIST λ^*, μ^* SUCH THAT WITH $\mathcal{L}(w, \lambda, \mu) = F(w) + \lambda^T G(w) + \mu^T H(w)$

$$\begin{aligned} \nabla_w \mathcal{L}(w^*, \lambda^*, \mu^*) &= 0 \\ G(w^*) &= 0 \\ 0 \leq \mu^* \perp H(w^*) &\leq 0 \end{aligned}$$

MUCH SIMPLER WITHOUT INEQUALITIES:

$$\begin{aligned} \nabla_w \mathcal{L}(w^*, \lambda^*) &= \nabla F(w^*) + \nabla G(w^*) \cdot \lambda^* = 0 \\ G(w^*) &\geq 0 \end{aligned} \quad \begin{matrix} n \\ m \end{matrix}$$

ONE CAN SOLVE THIS NONLINEAR SYSTEM IN ORDER TO FIND CANDIDATES $z = (w, \lambda)$ FOR PRIMAL & DUAL VARIABLES AT LOCAL MINIMIZERS. SUMMARIZE AS

$$\boxed{R(z) = 0} \quad \text{WITH } R(z) = \begin{bmatrix} \nabla_w \mathcal{L}(w, \lambda) \\ G(w) \end{bmatrix}$$

(SQUARE SYSTEM)

IDEA OF NEWTON-TYPE OPTIMIZATION:

GENERATE REFINED GUESSES $z_0, z_1, \dots$ BY SOLVING LINEAR SYSTEMS

$$z_{k+1} \in \left\{ z \mid \underbrace{R(z_k) + \frac{\partial R}{\partial z}(z_k)(z - z_k)}_{=: R_L(z; z_k)} = 0 \right\}$$

SOLUTION UNIQUE IF $\frac{\partial R}{\partial z}(z_n)$ INVERTIBLE:

(3)

$$z_{n+1} = z_n - \frac{\partial R}{\partial z}(z_n)^{-1} \cdot R(z_n)$$

STARTING AT z_0 , IT CAN DIVERGE OR CONVERGE.

IF IT CONVERGES TO A POINT z^* WITH $\frac{\partial R}{\partial z}(z^*)$ INVERTIBLE, IT CONVERGES Q-QUADRATICALLY

i.e.

$$\boxed{\|z_{n+1} - z^*\| \leq \frac{\omega}{2} \|z_n - z^*\|^2}$$

$$\text{IF } \left\| \frac{\partial R}{\partial z}(z)^{-1} \frac{\partial R}{\partial z}(z') - \mathbb{I} \right\| \leq \omega \|z - z'\| \quad \forall z, z' \in \mathbb{R}^{n+m}$$

(LIPSCHITZ-CONDITION ON $\frac{\partial R}{\partial z}(z)$)

PROOF:

$$\begin{aligned} z_{n+1} - z^* &= z_n - z^* - \frac{\partial R}{\partial z}(z_n)^{-1} (R(z_n) - R(z^*)) \\ &= \int_0^1 \left[\mathbb{I} - \frac{\partial R}{\partial z}(z_n)^{-1} \frac{\partial R}{\partial z}(z^* + t(z_n - z^*)) \right] dt \cdot (z_n - z^*) \end{aligned}$$

$$\|z_{n+1} - z^*\| \leq \int_0^1 \left\| \mathbb{I} - \frac{\partial R}{\partial z}(z_n)^{-1} \frac{\partial R}{\partial z}(z(t)) \right\| dt \cdot \|z_n - z^*\|$$

$z(t) = z_n + (1-t)(z^* - z_n)$

$$\leq \int_0^1 \omega \|z_n - z(t)\| dt \cdot \|z_n - z^*\|$$

$$\leq \int_0^1 \omega \|z_n - z^*\| (1-t) dt \cdot \|z_n - z^*\| = \frac{\omega}{2} \|z_n - z^*\|^2$$

□

(4)

IN OPTIMIZATION,

$$\frac{\partial R}{\partial z}(z) = \begin{bmatrix} \nabla_w^2 \mathcal{L}(w, \lambda) & \frac{\partial \mathcal{L}}{\partial w}(w)^T \\ \frac{\partial \mathcal{G}}{\partial w}(w) & 0 \end{bmatrix} \quad \text{"KKT matrix"}$$

EXACT

HESSIAN $\nabla_w^2 \mathcal{L}(w_k, \lambda_k)$ OFTEN APPROXIMATED BY B_k ,SOMETIMES ALSO JACOBIAN $\frac{\partial \mathcal{G}}{\partial w}(w)$ IS APPROXIMATED.→ "NEWTON-TYPE OPTIMIZATION"WITH SOME $M_k \approx \frac{\partial R}{\partial z}(z_k)$ • CONVERGENCEQ-LINEAR OBEYING

$$\|z_{k+1} - z^*\| \leq \left(\kappa + \frac{\omega}{2} \|z_k - z^*\| \right) \|z_k - z^*\|$$

IF $\|M_k^{-1}(M_k - \frac{\partial R}{\partial z}(z_k))\| \leq \kappa$ AND $\forall z_k, z \in \mathbb{R}^{n+m}$

$$\|M_k^{-1}(\frac{\partial R}{\partial z}(z) - \frac{\partial R}{\partial z}(z_k))\| \leq \omega \|z - z_k\| \quad (\text{NEARLY AS GOOD})$$

MOST IMPORTANT NEWTON-TYPE METHOD -

CONSTRAINED GAUSS-NEWTONFOR $F(w) = \frac{1}{2} \|r(w)\|_2^2$

WITH $M_k = \begin{bmatrix} B_k^{GN} & \frac{\partial \mathcal{G}}{\partial w}(w_k)^T \\ \frac{\partial \mathcal{G}}{\partial w}(w_k) & 0 \end{bmatrix}$

WITH "GAUSS-NEWTON HESSIAN"

$$B_k^{GN} := \frac{\partial r}{\partial w}(w_k)^T \frac{\partial r}{\partial w}(w_k)$$

WHERE

$$\kappa = O(\|r(w^*)\|)$$

TWO WAYS TO ADDRESS INEQUALITY CONST, NLP: ⑤

a) NONLINEAR INTERIOR POINT METHOD (IPM) EG. IPOPT.

b) SEQUENTIAL QUADRATIC PROGRAMMING (SQP)
EG. acados, uqp, filter-sqp.
(FROM KKT)

a) LIFTED-ROOT FINDING PROBLEM WITH BARRIER PARAM $\tilde{\gamma} > 0$

$$R_{\tilde{\gamma}}(z) = \begin{bmatrix} \nabla_w \mathcal{L}(w, \lambda, \mu) \\ G(w) \\ H(w) + S \\ S_1 \mu_1 - \tilde{\gamma} \\ \vdots \\ S_p \mu_p - \tilde{\gamma} \end{bmatrix}$$

$$z = (w, S, \lambda, \mu)$$

ENSURE $S > 0, \mu > 0$
($S > 0, \mu > 0$)

SOLVE FOR DESCENDING $\tilde{\gamma}$

b) AT GUESS w_k WITH HESSIAN $B_k \approx \nabla_w^2 \mathcal{L}(w_k, \lambda_k, \mu_k)$
SOLVE QUADRATIC PROGRAM (QP) TO GET NEW GUESS

$$w_{k+1} \in \arg \min_w F_L(w; w_k) + \frac{1}{2} (w - w_k)^T B_k (w - w_k)$$

$$\text{s.t. } G_L(w; w_k) = 0$$

$$H_L(w; w_k) \leq 0$$

WITH "LINEARIZATIONS AT w_k ".

MIKE POWELL: "IT CAN BE PROGRAMMED IN AN AFTERNOON IF ONE HAS A QP-SOLVER AVAILABLE"

(NOTE: QP-SOLVER IS OFTEN AN IPM)
INSIDE SQP

"LIFTED" NEWTON METHOD

6

MOST NONLINEAR FUNCTIONS CONTAIN "INTERMEDIATE VARIABLES".

THESE CAN BE "LIFTED" TO BECOME VISIBLE TO THE NEWTON-TYPE SOLVER, OFTEN WITH BETTER QUADRATIC CONVERGENCE SPEED (I.E. w SMALLER)

ROOT-FINDING EXAMPLE:

$$R(u) := u^4 - 2 \quad \text{AIM: SOLVE } \boxed{R(u) = 0}$$

(UNLIFTED)

EQUIVALENT LIFTED SYSTEM IN $w = (u, x)$

$$G(u, x) := \begin{bmatrix} u^2 - x \\ x^2 - 2 \end{bmatrix}$$

$$\boxed{G(u, x) = 0}$$

(LIFTED SYSTEM)

INTERESTING FACT: CAN CREATE AND SOLVE LIFTED LINEAR SYSTEMS

$$\boxed{G(w_h) + \frac{\partial G}{\partial w}(w_h)(w - w_h) = 0}$$

AT (NEARLY) SAME COST AS UNLIFTED LIN. SYSTEM

cf. Aldesmeijer & Diehl 2010, SIOP.

LIFTED CONSTRAINED GAUSS-NEWTON IS PARTICULARLY EXCITING...