Sequential Convex Quadratic Programming

R. Verschueren, N. van Duijkeren, R. Quirynen and M. Diehl

Syscop group retreat
5-7 September 2016

The Generalized Gauss-Newton algorithm

Consider a constrained nonlinear least-squares problem:

$$\begin{aligned} \underset{w \in \mathbb{R}^n}{\text{minimize}} \quad & \psi_0(w) = \frac{1}{2}\|c_0(w)\|_2^2 \\ \text{subject to} \quad & g(w) = 0 \\ & \psi(w) \leq 0. \end{aligned}$$

GGN algorithm

```
Find initial guess w_0.
for i=0,1,2,... do
    if converged then
        exit
    w_{i+1} = w_i + Δw
end for
```

$$\begin{aligned} \Delta w = \underset{\Delta w \in \mathbb{R}^n}{\arg\min} \quad & \frac{1}{2}\left\|c_0(w_i) + \frac{\partial c_0}{\partial w}(w_i)\Delta w\right\|_2^2 \\ \text{subject to} \quad & g(w_i) + \frac{\partial g}{\partial w}(w_i)\Delta w = 0 \\ & \psi(w_i) + \frac{\partial \psi}{\partial w}(w_i)\Delta w \leq 0. \end{aligned}$$

The Generalized Gauss-Newton algorithm

Consider a constrained nonlinear least-squares problem:

$$\begin{aligned} \underset{w \in \mathbb{R}^n}{\text{minimize}} \quad & \psi_0(w) = \frac{1}{2}\|c_0(w)\|_2^2 \\ \text{subject to} \quad & g(w) = 0 \\ & \psi(w) \leq 0. \end{aligned}$$

GGN algorithm

Find initial guess w_0.
for i=0,1,2,... **do**
    if converged **then**
        exit
    $w_{i+1} = w_i + \Delta w$
end for

$$\begin{aligned} \Delta w = \underset{\Delta w \in \mathbb{R}^n}{\arg\min} \quad & \frac{1}{2}\left\|c_0(w_i) + \frac{\partial c_0}{\partial w}(w_i)\Delta w\right\|_2^2 \\ \text{subject to} \quad & g(w_i) + \frac{\partial g}{\partial w}(w_i)\Delta w = 0 \\ & \psi(w_i) + \frac{\partial \psi}{\partial w}(w_i)\Delta w \leq 0. \end{aligned}$$

The Generalized Gauss-Newton algorithm

Consider a constrained nonlinear least-squares problem:

$$\begin{aligned} \underset{w \in \mathbb{R}^n}{\text{minimize}} \quad & \psi_0(w) = \frac{1}{2} \| c_0(w) \|_2^2 \\ \text{subject to} \quad & g(w) = 0 \\ & \psi(w) \leq 0. \end{aligned}$$

GGN algorithm

```
Find initial guess w_0.
for i=0,1,2,... do
    if converged then
        exit
    w_{i+1} = w_i + Δw
end for
```

$$\begin{aligned} \Delta w = \underset{\Delta w \in \mathbb{R}^n}{\arg\min} \quad & \frac{1}{2} \left\| c_0(w_i) + \frac{\partial c_0}{\partial w}(w_i) \Delta w \right\|_2^2 \\ \text{subject to} \quad & g(w_i) + \frac{\partial g}{\partial w}(w_i) \Delta w = 0 \\ & \psi(w_i) + \frac{\partial \psi}{\partial w}(w_i) \Delta w \leq 0. \end{aligned}$$

GGN and SQP

Different view on GGN:

full-step SQP method with Hessian approximation

$$B^{\mathrm{GN}} := \frac{\partial c_0}{\partial w}(w_i)^\top \frac{\partial c_0}{\partial w}(w_i).$$

Newton-type optimization and GGN

Two things you (maybe) did not know about Newton-type optimization:

- ▷ Necessary and sufficient condition for asymptotic stability
- ▷ Statistical stability (next group retreat)

A necessary and sufficient condition for asymptotic stability of a local minimizer

Consider the unconstrained problem,

$$w_{i+1} = w_i - B(w_i)^{-1} \nabla \psi_0(w_i).$$

Lemma (Linear Stability Analysis)

Regard iterations $w_{i+1} = F(w_i)$ with F a continuously differentiable function in a neighborhood of a fixed point $F(w^*) = w^*$.

$$\rho\left(\frac{\partial F}{\partial w}(w^*)\right) < 1 \quad \Longleftrightarrow \quad w^* \text{ is asymptotically stable.}$$

ρ is the spectral radius

A necessary and sufficient condition for asymptotic stability of a local minimizer

Consider the unconstrained problem,

$$w_{i+1} = w_i - B(w_i)^{-1}\nabla\psi_0(w_i).$$

Lemma (Linear Stability Analysis)

Regard iterations $w_{i+1} = F(w_i)$ with F a continuously differentiable function in a neighborhood of a fixed point $F(w^*) = w^*$.

$$\rho\left(\frac{\partial F}{\partial w}(w^*)\right) < 1 \quad \Longleftrightarrow \quad w^* \textit{ is asymptotically stable.}$$

ρ is the spectral radius

A necessary and sufficient condition for asymptotic stability of a local minimizer

Theorem (Bounds on Hessian approximation, unconstrained case)

Local minimizer $w^\star$ is asymptotically stable with asymptotic contraction rate $0 \leq \alpha < 1$, if and only if the following conditions hold:

$$\frac{\nabla^2 \psi_0(w^\star)}{1+\alpha} \preceq B(w^\star) \preceq \frac{\nabla^2 \psi_0(w^\star)}{1-\alpha}.$$

This theorem also holds for constrained problems.

A necessary and sufficient condition for asymptotic stability of a local minimizer

Theorem (Bounds on Hessian approximation, unconstrained case)

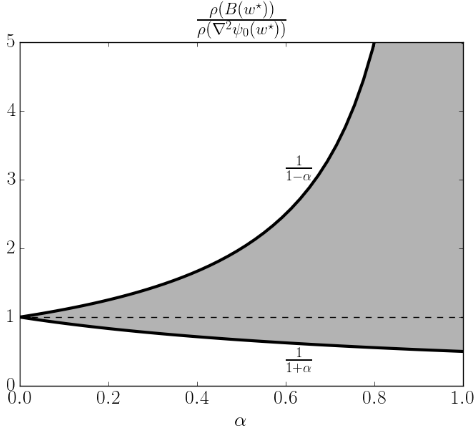
Local minimizer $w^\star$ is asymptotically stable with asymptotic contraction rate $0 \leq \alpha < 1$, if and only if the following conditions hold:

$$\frac{\nabla^2 \psi_0(w^\star)}{1+\alpha} \preceq B(w^\star) \preceq \frac{\nabla^2 \psi_0(w^\star)}{1-\alpha}.$$

This theorem also holds for constrained problems.

A necessary and sufficient condition for asymptotic stability of a local minimizer

$$\frac{\nabla^2 \psi_0(w^\star)}{1+\alpha} \preceq B(w^\star) \preceq \frac{\nabla^2 \psi_0(w^\star)}{1-\alpha}$$



Sequential Convex Quadratic Programming: a generalization of GGN

Consider

$$
\begin{aligned}
\min_{w \in \mathbb{R}^n} \quad & \phi_0(c_0(w)) && \text{(1a)}\\
\text{s.t.} \quad & g_i(w) = 0, \qquad i = 1, \dots, p, && \text{(1b)}\\
& \phi_i(c_i(w)) \leq 0, \; i = 1, \dots, q. && \text{(1c)}
\end{aligned}
$$

with $\phi_{0,1,\dots}(\cdot)$ convex.

$$
\begin{aligned}
B^{\mathrm{SCQP}}(w,\mu) := & \frac{\partial c_0}{\partial w}(w)^\top \nabla_c^2 \phi_0(c_0(w)) \frac{\partial c_0}{\partial w}(w) \\
& + \sum_{i=1}^{q} \mu_i \frac{\partial c_i}{\partial w}(w)^\top \nabla_c^2 \phi_i(c_i(w)) \frac{\partial c_i}{\partial w}(w).
\end{aligned}
$$

Sequential Convex Quadratic Programming: a generalization of GGN

Consider

$$\min_{w\in\mathbb{R}^n} \quad \phi_0(c_0(w)) \tag{1a}$$

$$\text{s.t.} \quad g_i(w) = 0, \qquad i = 1, \ldots, p, \tag{1b}$$

$$\phi_i(c_i(w)) \leq 0, \; i = 1, \ldots, q. \tag{1c}$$

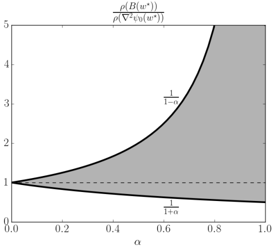
with $\phi_{0,1,\ldots}(\cdot)$ convex.

$$B^{\mathrm{SCQP}}(w,\mu) := \frac{\partial c_0}{\partial w}(w)^\top \nabla_c^2 \phi_0(c_0(w)) \frac{\partial c_0}{\partial w}(w) + \sum_{i=1}^{q} \mu_i \frac{\partial c_i}{\partial w}(w)^\top \nabla_c^2 \phi_i(c_i(w)) \frac{\partial c_i}{\partial w}(w).$$

Advantages

- ▷ SCQP is convex
- ▷ SCQP as cheap as GGN
- ▷ better approximation of exact Hessian

$\Longrightarrow$ Ideal for embedded optimization!



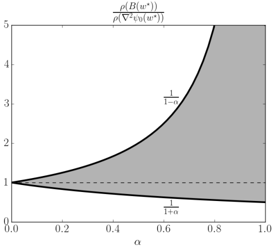
Advantages

▷ SCQP is convex

▷ SCQP as cheap as GGN

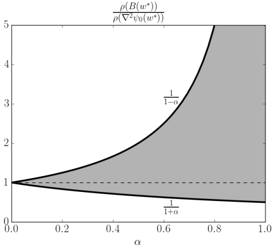
▷ better approximation of exact Hessian

$\Longrightarrow$ Ideal for embedded optimization!



Advantages

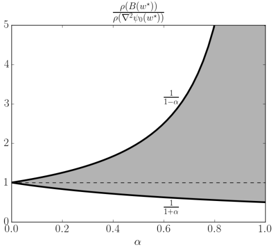
- ▷ SCQP is convex
- ▷ SCQP as cheap as GGN
- ▷ better approximation of exact Hessian



$\Longrightarrow$ Ideal for embedded optimization!

Advantages

- ▷ SCQP is convex
- ▷ SCQP as cheap as GGN
- ▷ better approximation of exact Hessian



$\Longrightarrow$ Ideal for embedded optimization!

Relation to SCP

In contrast to SCQP, SCP keeps nonlinear convex functions in constraints:

$$\begin{aligned} \underset{\Delta w \in \mathbb{R}^n}{\text{minimize}} \quad & f^\top (w_i + \Delta w) \\ \text{subject to} \quad & g(w_i) + \frac{\partial g}{\partial w}(w_i)\Delta w = 0 \\ & w_i + \Delta w \in \Omega, \end{aligned}$$

with Ω convex.

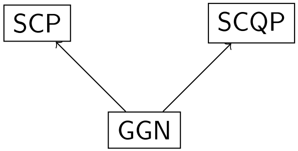
SCQP is an alternative to SCP as a generalization of GGN:

Relation to SCP

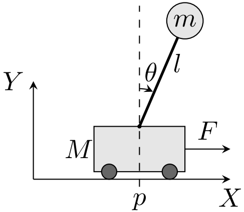
In contrast to SCQP, SCP keeps nonlinear convex functions in constraints:

$$\begin{aligned} \underset{\Delta w \in \mathbb{R}^n}{\text{minimize}} \quad & f^\top (w_i + \Delta w) \\ \text{subject to} \quad & g(w_i) + \frac{\partial g}{\partial w}(w_i)\Delta w = 0 \\ & w_i + \Delta w \in \Omega, \end{aligned}$$

with Ω convex.
SCQP is an alternative to SCP as a generalization of GGN:



Numerical example: inverted pendulum swing-up



$$\begin{bmatrix} X_{\text{mass}} \\ Y_{\text{mass}} \end{bmatrix} = \begin{bmatrix} p - l\sin(\theta) \\ l\cos(\theta) \end{bmatrix}$$

We solve the following OCP for different radii R_e of the terminal region:

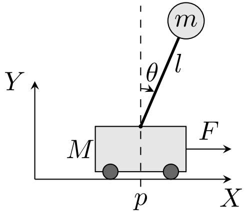
$$\min_{\substack{x_0,\ldots,x_N \\ u_0,\ldots,u_{N-1}}} \quad \frac{1}{2}\sum_{k=0}^{N-1} r_k F_k^2, \tag{2a}$$

$$\text{s.t.} \quad x_0 = \bar{x}_0, \tag{2b}$$

$$x_{k+1} = f(x_k, u_k), \quad k = 0, \ldots, N-1, \tag{2c}$$

$$\|[X_N - l, Y_N - l]^\top\|_2^2 - R_e^2 \leq 0, \tag{2d}$$

Numerical example: inverted pendulum swing-up



$$\begin{bmatrix} X_{\text{mass}} \\ Y_{\text{mass}} \end{bmatrix} = \begin{bmatrix} p - l\sin(\theta) \\ l\cos(\theta) \end{bmatrix}$$

We solve the following OCP for different radii R_e of the terminal region:

$$\min_{\substack{x_0,\dots,x_N \\ u_0,\dots,u_{N-1}}} \quad \frac{1}{2}\sum_{k=0}^{N-1} r_k F_k^2, \tag{2a}$$

$$\text{s.t.} \quad x_0 = \bar{x}_0, \tag{2b}$$

$$x_{k+1} = f(x_k, u_k), \quad k = 0, \dots, N-1, \tag{2c}$$

$$\|[X_N - l, Y_N - l]^\top\|_2^2 - R_e^2 \le 0, \tag{2d}$$

A closer look at the Hessians

GN Hessian

$$\begin{bmatrix} 0 & & & & & \\ & r_k & & & & \\ & & 0 & & & \\ & & & r_k & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

SCQP Hessian

$$\begin{bmatrix} 0 & & & & & \\ & r_k & & & & \\ & & 0 & & & \\ & & & r_k & & \\ & & & & \ddots & \\ & & & & & \nabla^2_{x_N} c_N(x_N) \end{bmatrix}$$

A closer look at the Hessians

GN Hessian

$$\begin{bmatrix} 0 & & & & & \\ & r_k & & & & \\ & & 0 & & & \\ & & & r_k & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

SCQP Hessian

$$\begin{bmatrix} 0 & & & & & \\ & r_k & & & & \\ & & 0 & & & \\ & & & r_k & & \\ & & & & \ddots & \\ & & & & & \nabla^2_{x_N} c_N(x_N) \end{bmatrix}$$

Numerical example: results

GGN does not converge..

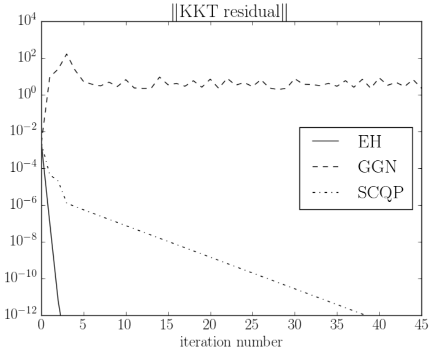


Figure: $R_e = 0.05\,\text{m}$.

Numerical example: results

GGN does not converge..

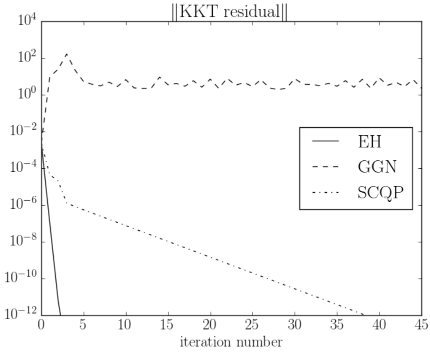


Figure: $R_e = 0.05\,\mathrm{m}$.

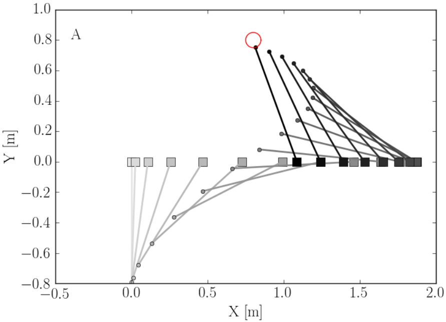


Figure: Trajectory of pendulum.

Numerical example: compare radii

GGN still does not converge!

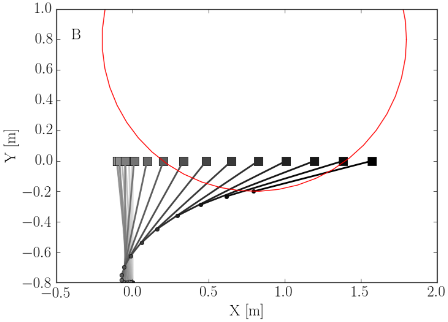


Figure: $R_e = 1\,\mathrm{m}$.

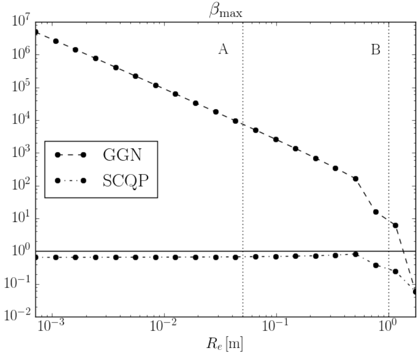


Figure: Comparison for different radii.

Numerical example: compare radii

GGN still does not converge!

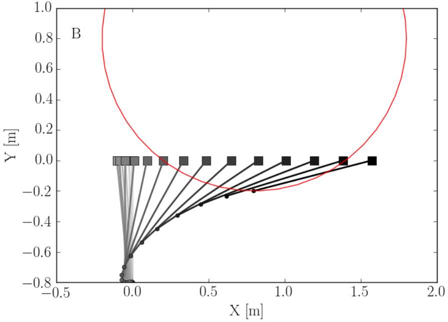


Figure: $R_e = 1\,\mathrm{m}$.

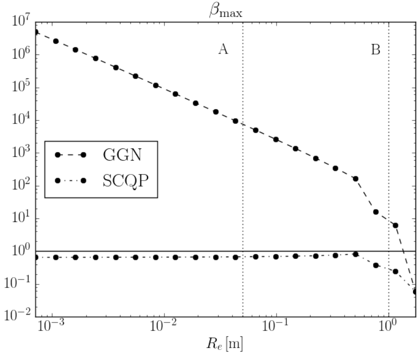


Figure: Comparison for different radii.

Conclusion

What we've done

- A new Hessian approximation for embedded SQP

What we want to do next

- Efficient implementation (acados!)

- Real-world tests

Conclusion

What we've done

- A new Hessian approximation for embedded SQP

What we want to do next

- Efficient implementation (`acados`!)

- Real-world tests

Thank you for your attention.
Questions?