Nonlinear Least Squares/Gauss-Newton

- What if the \( f \) to be minimized is actually a 2-norm?

\[
f(x) = \|r(x)\|_2, \quad r(x) = y - f(x)
\]

\[\hat{x}_0 = \ldots\]

\[\hat{x}_{n+1} = \hat{x}_n + \delta\]

**Demo:** Gauss-Newton

\[
\nabla \phi = \frac{\partial^2 \hat{r}(x)}{\partial x} \hat{r}(x)
\]

\[
H_\phi(x) = \frac{\partial^2 \hat{r}(x)}{\partial x} \hat{r}(x) + \sum \frac{\partial^2 \hat{r}(x)}{\partial r_i(x)} \frac{\partial \hat{r}(x)}{\partial r_i(x)} \Rightarrow H_\phi(\hat{x}_m) \delta = -\nabla \phi(\hat{x}_m)
\]

\[
\hat{H}_\phi(x) = \frac{\partial^2 \hat{r}(x)}{\partial x} \Rightarrow \hat{H}_\phi(\hat{x}_m) \delta = -\nabla \phi(\hat{x}_m)
\]

\[A\hat{x} = b\]

\[\hat{y}(\hat{x}) = \hat{r}(\hat{x})\]

\[\min \phi\]

\[\underbrace{\frac{1}{2} \| r(x) \|_2^2}_{\phi(x)} \]
\[ \hat{H}_p(\hat{x}_w) \hat{\delta} = -\hat{\gamma}_p(\hat{x}_w) \]
\[ \hat{\delta}_w = -\hat{\gamma}_w(x) \quad \hat{\delta}_s = -\hat{\delta}(x) \]
\[ (\hat{\gamma}_f \hat{\delta}_w + H_u \delta_n) \delta = -\hat{\gamma}_w(x) \quad \text{possible to write down as equ.} \]
\[ \text{Levenberg-Marquardt} \]

\[ \]
6.4 Constrained Optimization
Constrained Optimization: Problem Setup

- Want \( x^* \) so that

\[
f(x^*) = \min_x f(x) \quad \text{subject to} \quad g(x) = 0
\]

No inequality constraints just yet. This is equality-constrained optimization. Develop a necessary condition for a minimum.

**Unconstrained**: \( \nabla f(\bar{x}) = 0 \)

**Expectation**: At min, \( f(\bar{x} + \bar{\alpha} \bar{z}) \geq f(\bar{x}) \). Feasible direction \( \bar{z} \)

\[
f(\bar{x}^*) + \nabla f(\bar{x}^*) \cdot \bar{z} \geq f(\bar{x})
\]

\( \nabla f(\bar{x}^*) \cdot \bar{z} > 0 \)

for some small \( \bar{\alpha} \) on the interior of the feasible set

\[
\nabla f(\bar{x}^*) \cdot \bar{z} \geq 0 \Rightarrow \nabla f(\bar{x}^*) = 0
\]
Necessary cond:
\[ \nabla \mathcal{L}(\hat{x}, \hat{\lambda}) = g(x) + \lambda_0 \nabla g(x) \]

\[ \nabla \mathcal{L}(\hat{x}, \hat{\lambda}) = \left( \frac{\partial \mathcal{L}}{\partial x} \right) = \left( \nabla f + \partial g \lambda \right) g(x) = 0 \]
Lagrange Multipliers

\[ \nabla f(x) \succeq 0 \]

\[ \nabla f \in \text{rowspan}(J_g) \]

\[ \exists \lambda : \nabla f = \nabla J_g \lambda \]

\[ \nabla g_i = \frac{\partial g_i}{\partial x_1, \ldots, x_n} \]

Seen: Need \[ \nabla g_i \neq 0 \]

\[ -\nabla f(x) = J_g^T \lambda \]

at the (constrained) optimum.

Idea: Turn constrained optimization problem for \( x \) into an unconstrained optimization problem for \( (x, \lambda) \). How?
Demo: Sequential Quadratic Programming
**Inequality-Constrained Optimization**

Want $x^*$ so that

$$f(x^*) = \min_x f(x) \text{ subject to } g(x) = 0 \text{ and } h(x) \leq 0$$

This is inequality-constrained optimization. Develop a necessary condition for a minimum.

Define Lagrangian:

$$\mathcal{L}(x, \lambda_1, \lambda_2) := f(x) + \lambda_1^T g(x) + \lambda_2^T h(x)$$

- Some inequality constrains may not be “active”
  (active $\iff h_i(x^*) = 0 \iff$ at ‘boundary’ of ineq. constraint)
  (Equality constrains are always ‘active’)
- If $h_i$ inactive ($h_i(x^*) < 0$), must force $\lambda_{2,i} = 0$. 