Nonlinear Least Squares/Gauss-Newton

$Ax=b$

$\vec{f}(\vec{x}) - \vec{y} = \vec{r}(\vec{x})$

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

$$f(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|_2, \qquad \boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x})$$

Demo: Gauss-Newton

$\vec{x}_0 = \ldots$

$\vec{x}_{k+1} = \vec{x}_k + \vec{s}$

$\varphi(x) = \frac{1}{2}\|\vec{r}(x)\|_2^2$

$\min \varphi$

$\nabla\varphi = J_{\vec{r}(x)}^T \vec{r}(x)$

$J = J_r(\vec{x})$

$H_\varphi(x) = J^T J + \sum_i r_i H_{r_i}(\vec{x}) \rightarrow H_\varphi(\vec{x}_k)\vec{s} = -\nabla_\varphi(\vec{x}_k)$

$\hat{H}_\varphi(x) = J^T J \rightarrow \tilde{H}_\varphi(\vec{x}_k)\tilde{\vec{s}} = -\nabla_\varphi(\vec{x}_k)$

Gauss-Newton

$$\tilde{H}_\varphi(\vec{x}_k)\,\tilde{\vec{s}} = -\nabla_\varphi(\vec{x}_k)$$

$$J^T J\,\tilde{\vec{s}} = -J^T \vec{r}(\vec{x}) \qquad J\tilde{\vec{s}} \tilde{=} -\vec{r}(\vec{x})$$

$$\left(J^T J\,\vec{s} + \mu_k 1\right)\vec{s} = -J^T \vec{r}(\vec{x}) \longrightarrow$$ possible to write down an equiv. least sq. prob.

Levenberg-Marquardt

6.4 Constrained Optimization

Constrained Optimization: Problem Setup

- Want $\boldsymbol{x}^*$ so that

$$f(\boldsymbol{x}^*) = \min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$$

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

Unconstrained: $\nabla f(\vec{x}) = \vec{0}$

Expectation: At min, $f(\vec{x}^* + \alpha\vec{s}) \ge f(\vec{x}^*)$.

feasible direction $\vec{s}$

$\vec{x} + \alpha\vec{s} \in S$

set of feasible points

$\cancel{f(\vec{x}^*)} + \nabla f(\vec{x}^*)\cdot\vec{s} \ge \cancel{f(\vec{x}^*)}$

$\nabla f(\vec{x}^*)\cdot\vec{s} > 0$

for some small α

On the interior of the feasible set

$x + \alpha\vec{s} \in S$, $x - \alpha\vec{s} \in S$,

$\nabla f(\vec{x}^*)\cdot\vec{s} \ge 0$

$\nabla f(\vec{x}^*)\cdot(-\vec{s}) \ge 0 \Rightarrow \nabla f(\vec{x}) = \vec{0}$

Necessary cond:

$$\nabla f(\vec{x}) + Jg^T \vec{\lambda}$$

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

$$\nabla \mathcal{L}(\vec{x}, \vec{\lambda}) = \begin{pmatrix} \nabla_x \mathcal{L} \\ \nabla_\lambda \end{pmatrix} = \begin{pmatrix} \nabla f + Jg^T \lambda \\ g(x) \end{pmatrix} = 0$$

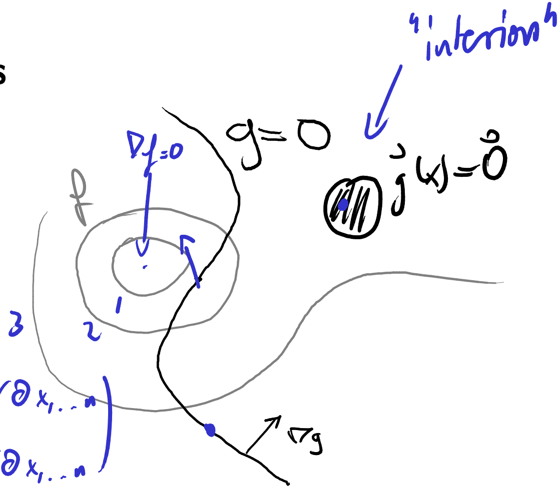
Lagrange Multipliers

$\nabla f(\vec{y}) \cdot \vec{s} \geq 0$

$\nabla f \in \mathrm{rowspan}(J_g)$

$\exists \vec{\lambda}: -\nabla f = J^T \vec{\lambda}$

$J_g = \begin{pmatrix} \partial g_1 / \partial x_{1 \ldots n} \\ \partial g_n / \partial x_{1 \ldots n} \end{pmatrix}$



- Seen: Need $\nabla g_i \cdot \nabla f \neq 0$

$$-\nabla f(\boldsymbol{x}) = J_{\boldsymbol{g}}^T \boldsymbol{\lambda}$$

at the (constrained) optimum.
Idea: Turn constrained optimization problem for $\boldsymbol{x}$ into an *unconstrained* optimization problem for $(\boldsymbol{x}, \boldsymbol{\lambda})$. How?

Demo: Sequential Quadratic Programming

Inequality-Constrained Optimization

Want $\boldsymbol{x}^*$ so that

$$f(\boldsymbol{x}^*) = \min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{g}(\boldsymbol{x}) = \mathbf{0} \quad \text{and} \quad \boldsymbol{h}(\boldsymbol{x}) \leqslant \mathbf{0}$$

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

Define Lagrangian:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) := f(\boldsymbol{x}) + \boldsymbol{\lambda}_1^T \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\lambda}_2^T \boldsymbol{h}(\boldsymbol{x})$$

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