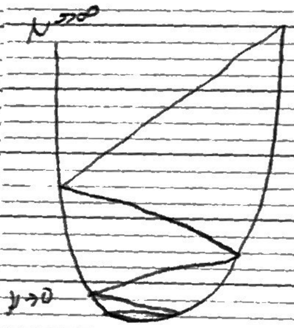
Levenberg - Marquardt method

Algo: Starting from an initial point, Repeat it

↳ Approx the hessian matrix: $H(\theta) \approx J(\theta)^T J(\theta) - \mu I_n$, $\mu > 0$

↳ compute descent direction by solving

$$[J(\theta^k)^T J(\theta^k) + \mu I_n] d^k = -J(\theta^k)^T F(\theta^k)$$



↳ Determine approximately λ^* w/ line search / golden search / ...

↳ Update the point $\theta^{k+1} = \theta^k + \lambda^* d^k$

Until the stopping conditions are satisfied

Remark:

↳ if $\mu = 0$; we get the gauss-newton method

↳ if $\mu \to \infty$; gradient method (dynamic adjustment)

↳ There are automatic ways to determin μ

↳ it is modified each iteration, starting very large (gradient) and moving towards 0 (gauss-newton) near the solution where the residual becomes small

---

Optimality Conditions

consider the problem w/ equality constraints only:

$\min_{\theta \in \mathbb{R}^n} J(\theta)$ s.t. $g_i(\theta) = 0$, $i = 1 \dots p$

Define the Lagrangian function

$$L(\theta, \lambda) = J(\theta) + \lambda^T g(\theta) \qquad \theta \in \mathbb{R}^n, \lambda \in \mathbb{R}^n$$

$$L(\theta, \lambda) = J(\theta) + \sum_{i=1}^{p} \lambda_i g_i(\theta)$$

where λ is the vector of Lagrange multipliers

The optimality conditions are given by the Lagrangian function: the optimal solution minimizes this function

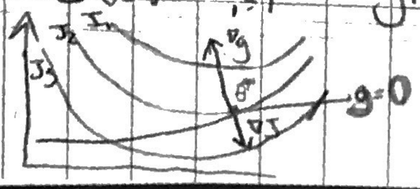
If θ^* is a local feasible minimizer:

$$\exists \lambda^* \in \mathbb{R}^n, \nabla L(\theta^*, \lambda^*) = 0$$

Since we have turned a constrained problem into an unconstrained problem where we want to minimize $L(\theta$

rewritten:

$$\nabla J(\theta^*) + \sum_{i=1}^{p} \lambda_i^* \nabla g_i(\theta^*) = 0 \quad , \text{ s.t. } g(\theta) = 0$$



Consider the problem w/ inequality constraints

$\min_{\theta} J(\theta)$ s.t. $h_i(\theta) \leq 0$, $i, \dots, m$

Define the Lagrangian function

$L(\theta, \mu) = J(\theta) + \mu^T h(\theta)$ $\quad \theta \in \mathbb{R}^n$ $\mu \in \mathbb{R}^{+m}$

where μ is the vector of Lagrange multipliers

If θ^* is a local feasible minimizer,

$\nabla J(\theta^*) + \sum_{i=1}^{m} \mu_i^* \nabla h_i(\theta) = 0$ s.t $h(\theta^*) = 0$

$\forall i$, $1 \leq i \leq m$ $\quad \mu_i^* h(\theta^*) = 0$

where $\mu_i \geq 0$, because from the optimal solution

we know that $\mu_i < 0$, we are moving away

If $\mu_i^* > 0 \Rightarrow h_i(\theta^*) = 0$ active set

If $h_i(\theta^*) < 0 \Rightarrow \mu_i^* = 0$ inactive set

$\mu^* = 0$, $\nabla J = 0$

$h>0$ $\quad h<0$ $\quad h=0$

$h = 0$, $\mu > 0$

$\nabla J = \sum \mu \nabla h_i(\theta)$

$J(x, y) = x^2 - xy + 5y^2 - 10$ $\quad 4.99$

$1 - 1 + 5 - 10 = -5$

$4 - 4 + 20 - 10 = 10$

2,2 $\quad 4 - 4 + 5 - 10 = -3$

2,1 $\quad 4 - 2 + 5 - 10$

$xy - 2(x - y - 4) = 0$ $\quad -x + y \pm 8$

$(y, x) = 2(-y, x)$