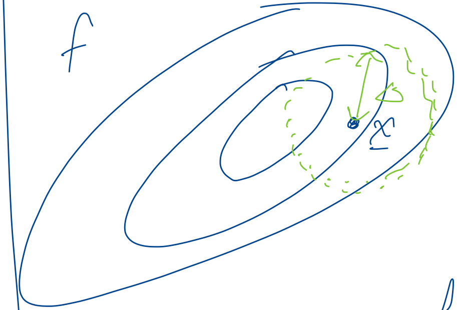
Trust Region Methods

- Trust region subproblem
- Idealized TR method
- Solving the TR subproblem
- The DOGLEG method
- The Cauchy point
- A global convergence result

Trust Region Methods

High level idea

- pick step length
- then pick direction.



Line Search: pick $\underline{p}$, then $\underline{x}^{+} = \underline{x} + \alpha \underline{p}$

Trust Region: pick Δ, then find $\underline{x}^{+} = \underline{x} + \underline{p}$, s.t. $\|p\| \leq$

Formally, the
that

$f(\underline{x}+\underline{p}) = f(\underline{x}$

$\bar{H}$ is a nearby

Idea: use B

then minimize

$M_K(p) =$

s.t. $\|\underline{p}\|$

idea is

$) + p^T g + \frac{1}{2} p^T \bar{H} p ,$

Hessian.

$\simeq \bar{H}$

$f(x_k) + p^T g_k + \frac{1}{2} p^T B_k p$

$\le \Delta .$

$\le \Delta$

If Δ is small enough
Then...

$$M_k(p) \simeq f(x_k + p)$$

The Trust Region Problem as Subproblem

$$\min_p \; M_k(p) = p^T g + \frac{1}{2} p^T B_k p + f(x_k)$$

$$\text{s.t.} \quad \|p\|_2 \le \Delta .$$

Choices for B_k

- Hessian at x_k
- QN Surrogate.
- Identity ??

Suppose $B_K \succ 0$.

If Δ is large enough,
What is p?

$$p = -B_K^{-1} g$$

if $\Delta \geq \|B_K^{-1} g\|$.

If $\Delta = \varepsilon$ (tiny)
then $p \approx -g/\|g\| \cdot \Delta$.

Idealized TR Method

- Given Δ_{max}
- Given $0 < \Delta_0 < \Delta_{max}$.

While not done

$$p = \operatorname{argmin} M_K(p) \quad \text{s.t.} \quad \|p\| \leq \Delta_K.$$

Let $\eta = \dfrac{f(x_k) - f(x_k + p)}{M_K(0) - M_K(p)}$

- If $\eta \approx 1$, keep Δ, take step.
- If $\eta \ll 1$, reduce Δ, resolve.
- If $\eta \gg 1$, increase Δ, resolve.

Don't need to solve Exactly!

More GD like

More Newton-like.

But How do we Solve the TR problem?

$$\min \; m_k(p) \text{ s.t. } \|p\| \le \Delta.$$

1. 1 ineq. constraint $\Rightarrow$ 1 Lagrange mult λ.

$\Leftrightarrow$

$$\lambda \Delta - \lambda \|p\| = 0.$$

Also, Lagrangian gives

$$(B_k + \lambda I) p = -g.$$

And $(B_k + \lambda I) \succeq 0$.

$p(\lambda) = -(B_k$

$\lambda \Delta - \lambda \|p(\lambda$

1d non-lin eq. Solve for λ

The code uses S

Binary Search +

Newton Steps to

this non-lin eq.

Still "too slow" i

practice.

$+ \lambda I)^{-1} g$

$\boxed{)\| = 0.}$

to

one

Bisection

Solve

The IDEA of the Dog leg method is to make this go faster.

$\Rightarrow$ this will give us the Cauchy pt.

Idea 0: Look at TR problem only in the neg. gradient dir.

$$\min \; f_k + p^T g \quad \text{s.t.} \quad \|p\| \le \Delta.$$

$$\Rightarrow p = -\frac{g_k}{\|g_k\|} \cdot \Delta.$$

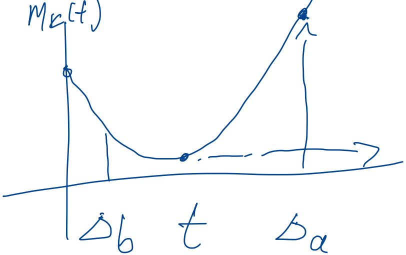
Idea 1: (True Cauchy pt.)

$$\min \; m_k(p) = f_k + p^T g + \tfrac{1}{2} p^T B_k p$$

$$\text{s.t.} \quad p = -t \, g_k / \|g_k\|$$

$$0 \le t \le \Delta.$$

What's the Diff?



Idea 1 defines the Cauchy pt.

$\underline{C}$:= Cauchy pt $\Leftrightarrow$ Soln of

Is it "Easy"? to solve for this point!

$\boxed{\text{Quiz}}$

Prob is Equiv to

$$\underline{d} = g/\|g\|$$

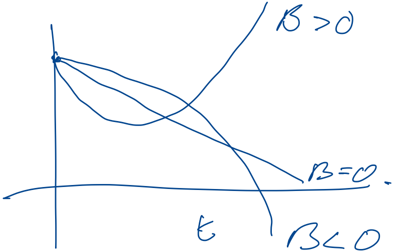
$$\min_t \; f_k - t\underbrace{(g^T d)}_{\alpha} + \frac{t^2}{2}\underbrace{(d^T B_k d)}_{\beta}$$

st $0 \le t \le \Delta$.

$$\min_t \; f_k - \alpha t + \frac{1}{2}\beta t^2 \Rightarrow$$

s.t. $0 \le t \le \Delta$.

3 Cases: $\beta > 0$, $\beta \le 0$.



If $\beta \le 0$,
$t = \Delta$.

If $\beta > 0$,
Compute m
of quadratic
$t_m = \dfrac{\alpha}{\beta}$.

If $\Delta < \dfrac{\alpha}{\beta}$,
$t = \Delta$

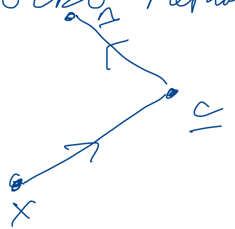
Else
$t = t_m$.

Set

$-\alpha + \beta t = 0$

$t = \frac{\alpha}{\beta}$

DOGLEG Method



find a parameterize curve to move from $x \to c$, then $c \to n$.

Theorem:

Given a

Method

$M_k(0) -$

$\geq C_1$

$0 \leq C$

Then we get a stationa

Using p_k^C

$C_1 =$

ny Trust Region

Where

$- M_k(p_k)$

$\|g_k\| \cdot \min\left(\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right)$

(

Global Convergence to

y pt.

(Cauchy pt). gives

$= 1/2$.

So any

$M_k(p_k) \le M_k(p_k^C)$

will be Globally Convergent.