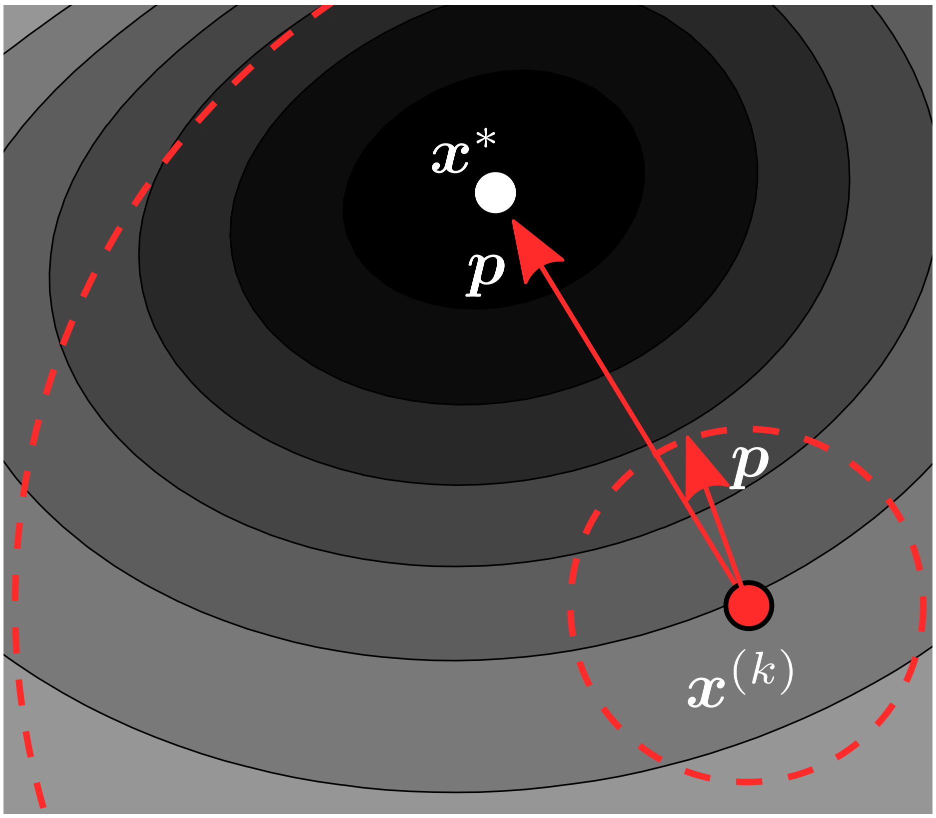
TRUST REGION METHODS

David F. Gleich

April 21, 2026

The material here is from Chapter 4 in Nocedal and Wright.

1 PICTORIAL DESCRIPTION OF TRUST REGION METHODS



2 THE TRUST REGION MODEL

Taylor's series gives us:

$$f(\mathbf{x}_k + \mathbf{p}) = f_k + \mathbf{g}_k^T \mathbf{p} + \frac{1}{2}\mathbf{p}^T \bar{\boldsymbol{H}} \mathbf{p}$$

where $\bar{\boldsymbol{H}} = \boldsymbol{H}(\mathbf{x}_k + t\mathbf{p})$ for some $0 \le t \le 1$. The ideal with the trust region method is to use this model to find a direction $\mathbf{p}$ within a "trusted region". That is, we know that this approximation is only valid in a small area, so let's pick something to approximate $\boldsymbol{B}_k \approx \bar{\boldsymbol{H}}$ and then find:

$$\begin{array}{ll} \underset{\mathbf{p}}{\text{minimize}} & f_k + \mathbf{g}_k^T \mathbf{p} + \frac{1}{2}\mathbf{p}^T \boldsymbol{B}_k \mathbf{p} \\ \text{subject to} & \|\mathbf{p}\| \le \Delta. \end{array}$$

3 FIRST OBSERVATIONS

If Δ is sufficiently large and $\boldsymbol{B}_k > 0$, then this will correspond to

$$\underset{\mathbf{p}}{\text{minimize}} \quad f_k + \mathbf{g}_k^T \mathbf{p} + \frac{1}{2}\mathbf{p}^T \boldsymbol{B}_k \mathbf{p} \quad .$$

The minimizer of this is $\boldsymbol{B}_k \mathbf{p} = -\mathbf{g}$, *the Newton step*. In trust region lingo, this is called *the full step*.

If Δ is sufficiently small, then what happens? We need a bit more info to tell us.

We'll work through this in a little bit, but the theory of constrained optimization tells us that the solution of this problem is:

$$(\boldsymbol{B}_k + \lambda \boldsymbol{I})\mathbf{p} = -\mathbf{g}.$$

where $\lambda\Delta - \lambda\|\mathbf{p}\| = 0$, and $\boldsymbol{B}_k + \lambda \boldsymbol{I}$ is positive semi-definite.

CLASS QUESTION What happens to $\|\mathbf{p}\|$ as λ gets larger?

This is related to Tikinov regularization! Recall that Tikinov regularization is:

$$\text{minimize} \quad \|\mathbf{b} - \boldsymbol{A}\mathbf{x}\|^2 + \lambda\|\mathbf{x}\|^2.$$

We could solve this system by solving:

$$(\boldsymbol{A}^T\boldsymbol{A} + \lambda \boldsymbol{I})\mathbf{x} = \boldsymbol{A}^T\mathbf{b}.$$

What did we see for $\lambda \to 0$ and $\lambda \to \infty$?

QUIZ What is the smallest value that λ could possibly be for a given $\boldsymbol{B}_k$? (debugging code motivation)

SOLUTION Since $\boldsymbol{B}_k + \lambda \boldsymbol{I}$ must be positive semi-definite, all its eigenvalues must be larger than 0. Let the eigenvalues of $\boldsymbol{B}_k$ be $\lambda_1 \leq \ldots \leq \lambda_n$. Then the eigenvalues of $\boldsymbol{B}_k + \lambda \boldsymbol{I}$ are $\lambda_1 + \lambda \leq \ldots \leq \lambda_n + \lambda$; we need $\lambda_1 + \lambda \geq 0$, so $\lambda \geq -\lambda_1$. Thus, solutions of the problem must have $\lambda \in [-\lambda_1, \infty)$.

If $\lambda = 0$, then we get the Newton step! So what happens as $\lambda \to \infty$? Obviously, $\mathbf{p} \to 0$.

QUIZ But what direction does it go to (if any) as $\lambda \to \infty$? Argue or prove. For large enough λ, the matrix $\boldsymbol{B}_k + \lambda \boldsymbol{I} \approx \lambda \boldsymbol{I}$, in which case, $\mathbf{p} \approx -\frac{1}{\lambda}\mathbf{g}$. This is the steepest descent direction.

Consequently, the trust region method "interpolates" between Newton's method (large Δ) and steepest descent (small Δ)

4 A TRUST REGION METHOD

```
Given a maximum trust region size Δ_max > 0
Given an initial region 0 < Δ_0 < Δ_max
Given a parameter η

while not done
  Compute the next step p_k by solving
      minimize_p   f_k + g_k^T p + (1/2) p^T B_k p      subject to ||p|| ≤ Δ_k  .

  Check how well your model matches the function by
  evaluating ρ_k = (f(x_k) − f(x_{k+1})) / (m_k(0) − m_k(p_k))

  # Update the trust region
  If ρ_k < 1/4, then set Δ_{k+1} = Δ/4
  If ρ_k > 3/4 and ||p_k|| = Δ_k, set Δ_{k+1} = min(2Δ, Δ_max)

  If ρ_k > η, x_{k+1} = x_k + p_k
  Else, x_{k+1} = x_k
```

5 SOLVING THE SUBPROBLEM

See `solve_trust_regions_subproblem.m`. It's nasty. It involves trying to find a root of $\|\mathbf{p}(\lambda)\| - \Delta$. Each iteration involves solving a linear system!

6 MAKING IT PRACTICAL

We need something faster! For line search methods, we didn't need an exact line search. Instead, we just needed to make enough progress. Something similar will hold true here.