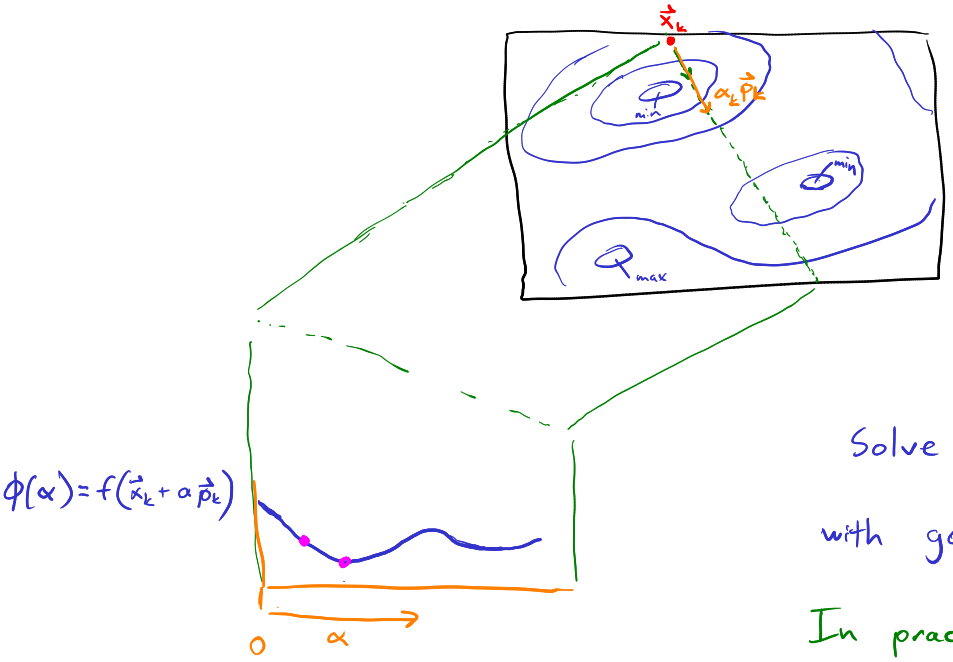
Today
Line Search
Newton's Method
Conjugate Gradient

Announcements
Exam 2 Monday
Lab 3 due tomorrow

Fixed step size gradient descent is bad!



$$\vec{p}_k \leftarrow -\nabla f(\vec{x}_k)$$

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

Solve $\underset{\alpha}{\text{minimize}}\ \phi(\alpha)$

with golden section search, bisection, etc.

In practice, exact line search is usually too expensive.

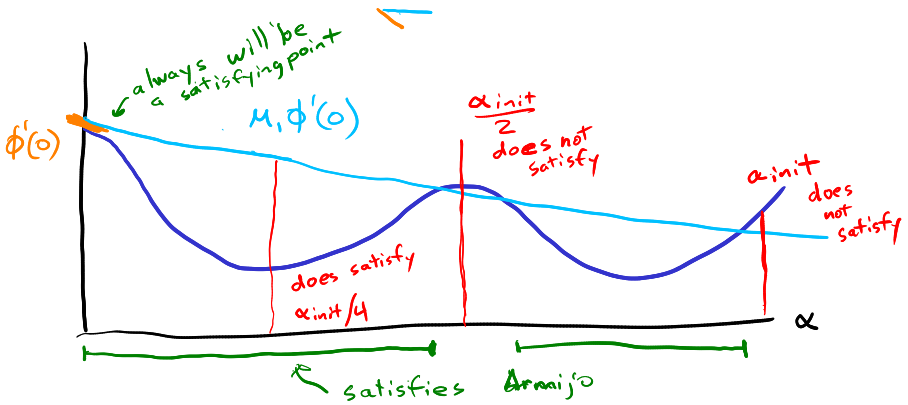
Better to find a "good enough" point ← sufficient decrease

Simple option: Armijo backtracking

Armijo condition: $\phi(\alpha) \le \phi(0) + \mu_1 \alpha \phi'(0)$

(μ_1: tuning param: usually 10^{-4})

($\phi'(0) = \nabla f(\vec{x}_k)^T \vec{p}_k$)

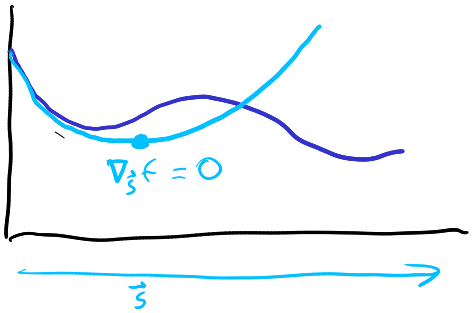


Armijo Backtracking Line Search

input : current point $\vec{x}_k$, descent direction $\vec{p}_k$, initial step α_{init} (1 for Newton), sufficient decrease factor μ_1 (e.g. 10^{-4}), backtracking factor ρ (e.g. 0.5)
output: step size α_k

```
φ(α) ← f(x_k + α p_k)                    // f along the line
φ'(0) ← ∇f(x_k)^T p_k                    // slope at the start, < 0
α ← α_init
while φ(α) > φ(0) + μ1 α φ'(0) do
    α ← ρ α                              // shrink the step
end
return α_k = α
```

Newton's Method



Taylor Series

$$f(\vec{x}_k + \vec{s}) = f(\vec{x}_k) + \nabla f(\vec{x}_k)^T \vec{s} + \frac{1}{2} \vec{s}^T H(\vec{x}_k) \vec{s} + \dots$$

$$\nabla_{\vec{s}} f(\vec{x}_k + \vec{s}) \approx \nabla f(\vec{x}_k) + H(\vec{x}_k) \vec{s}_k = 0$$

$$\vec{s} = -H(\vec{x}_k)^{-1} \nabla f(\vec{x}_k)$$

Newton update

$$\vec{x}_{k+1} \leftarrow \vec{x}_k - H(\vec{x}_k) \backslash \nabla f(\vec{x}_k)$$

Newton's Method

input : starting point $\vec{x}_0$, tolerance E_s
output: $\vec{x}^*$, estimate of a local minimizer

$k \leftarrow 0$
while $\|\nabla f(\vec{x}_k)\| > E_s$ **do**
  $\vec{p}_k \leftarrow -H(\vec{x}_k)\backslash\nabla f(\vec{x}_k)$ `// stationary point of quadratic fit`
  if $\nabla f(\vec{x}_k)^\mathsf{T}\vec{p}_k \geq 0$ **then**
    $\vec{p}_k \leftarrow -\nabla f(\vec{x}_k)$ `// not downhill: use steepest descent`
  end
  $\alpha_k \leftarrow$ line search along $\vec{p}_k$ with $\alpha_{\text{init}} = 1$
  $\vec{x}_{k+1} \leftarrow \vec{x}_k + \alpha_k \vec{p}_k$
  $k \leftarrow k + 1$
end
return $\vec{x}^* = \vec{x}_k$

Newton's gives a better step size AND a better search direction.