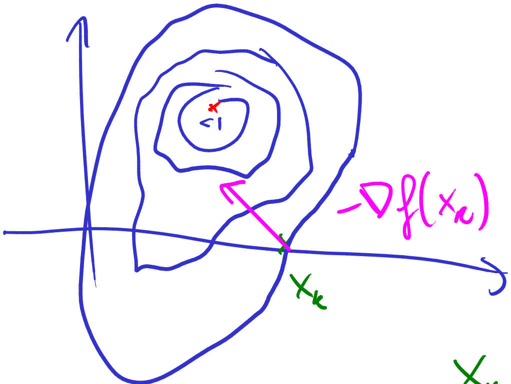
6.2 Methods for unconstrained opt. in *n* dimensions

$$f : \mathbb{R}^n \to \mathbb{R}$$



$$x_{k+1} = \underbrace{x_k + \alpha\left(-\nabla f(x_k)\right)}_{\varphi(\alpha)}$$

Steepest Descent

- Given a scalar function $f:\mathbb{R}^n \to \mathbb{R}$ at a point $\boldsymbol{x}$, which way is down?

1. $x_0 =$ (starting guess)
2. Define helper $\varphi(\alpha) := f(x_k + \alpha(-\nabla f(x_k)))$
3. Find min α' of φ
4. $x_{k+1} = x_k + \alpha(-\nabla f(x_k))$

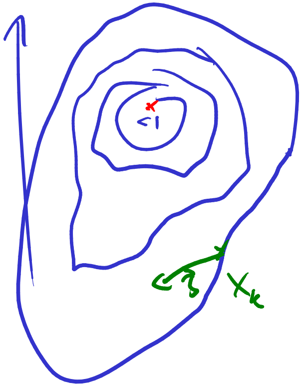
Demo: Steepest Descent

Newton's method (nD)

$$\nabla f(\vec{x}) \cdot \vec{s} = \nabla f(\vec{x})^T \vec{s} \qquad \vec{x} \cdot \vec{y} = \vec{x}^T y$$

- What does Newton's method look like in n dimensions?

$$f(\vec{x}+\vec{s}) \approx f(\vec{x}) + \underbrace{\nabla f(\vec{x}) \cdot \vec{s}} + \frac{1}{2}\vec{s}^T H_f(\vec{x})\vec{s} = \varphi(\vec{s})$$



$$\nabla_{\vec{s}}\varphi(\vec{s}) = \nabla f(\vec{x}) + H_f(\vec{x})\vec{s} \overset{!}{=} 0$$

$$H_f(\vec{x})\vec{s} = -\nabla f(x)$$

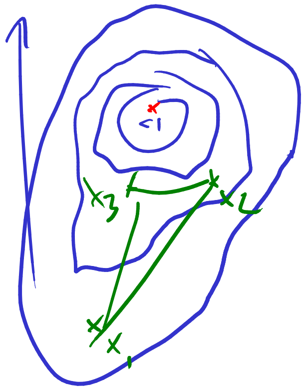
$$\vec{s} = -H_f(\vec{x})^{-1}\nabla f(\vec{x}) \qquad s = -\frac{f'(x)}{f''(x)}$$

x_0 = starting guess

$$x_{k+1} = x_k - H_f(\vec{x}_k)^{-1}\nabla f(\vec{x}_k)$$

Demo: Newton's method in n dimensions

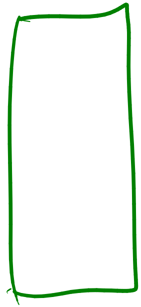
Demo: Nelder-Mead Method



6.3 Nonlinear Least Squares

$$Ax \cong b$$

$$\min \|Ax - b\|_2^2$$



$$f(x) = y$$

$$\min \|f(x) - y\|_2^2$$

Nonlinear Least Squares/Gauss-Newton

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

$$f(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|_2^2, \qquad \boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x})$$

Demo: Gauss-Newton

In-class activity: Optimization II

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

$$\frac{\partial}{\partial x_j}\varphi = \frac{\partial}{\partial x_j}\left[\frac{1}{2}\sum_i r_i(x)^2\right]$$

$$\partial_x f^2(x) = 2 \cdot f'(x) f(x)$$

$$= \sum_i \left[\frac{\partial}{\partial x_j} r_i(x)\right] r_i(x)$$

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

$$J = J_{\vec{r}}(\vec{x})$$

$$H_\varphi(x) = J^T J + \underbrace{\sum_i r_i H_{r_i}}_{\text{small if residual close to 0} \rightarrow \text{ignore outright}}$$

$x_0 = \langle$ starting guess $\rangle$

$$H_\varphi s = -\nabla\varphi$$

$$x_{k+1} = x_k - \underbrace{H_\varphi^{-1}\nabla\varphi}_{s}$$

$$J^T J s = -J^T r$$

$$J s \approx -r$$

$$A^T A x = A^T b$$

$$A x \approx b$$

6.4 Constrained Optimization

Constrained Optimization: Problem Setup

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

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

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