Orders
- No derivatives — 0th order
- 1st derivatives, — 1st order (calculate Jf)
- 2nd derivaties — 2nd order (calculate Hessian)

$$f(x_1, x_2, x_3, x_4) \longrightarrow S$$

$$f(x) \longrightarrow S$$

Numerical derivatives

$$f'(x) = \frac{f(x+h) - f(x)}{h}$$

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

$$f''(x) = \frac{\frac{f(x+h) - f(x)}{h} - \frac{f(x) - f(x-h)}{h}}{h}$$

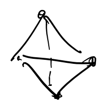
$$= \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

central difference

Zeroth order

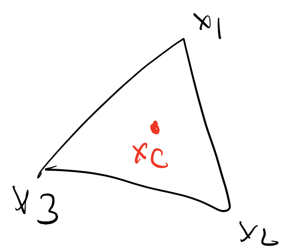
Nelder-Mead → Simplex is $n+1$ vertices in n dimension

$$\left.\begin{array}{l} \text{1D} \rightarrow \text{line segment} \\ \text{2D} \rightarrow \triangle \ \text{triangle} \\ \text{3D} \rightarrow \text{tetrahedron} \end{array}\right\} \text{simplex}$$



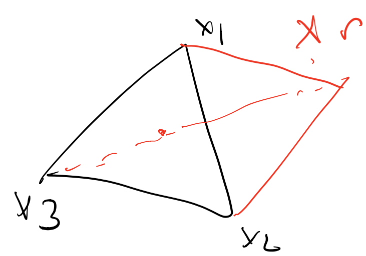
* Heuristic but works well in practice & robust

① initial simplex

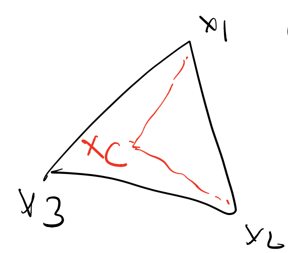


$$f(x_1) \le f(x_2) \le f(x_3)$$

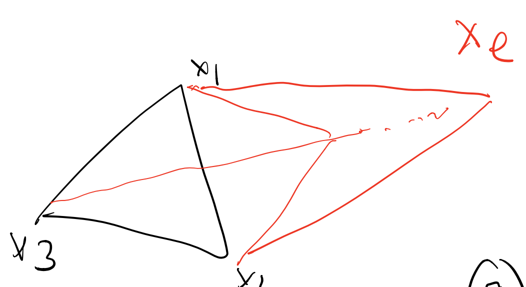
② reflect

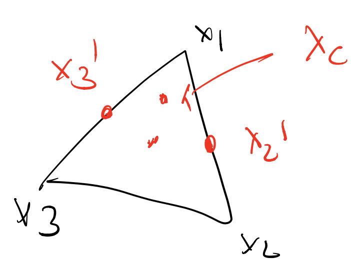


④ contract



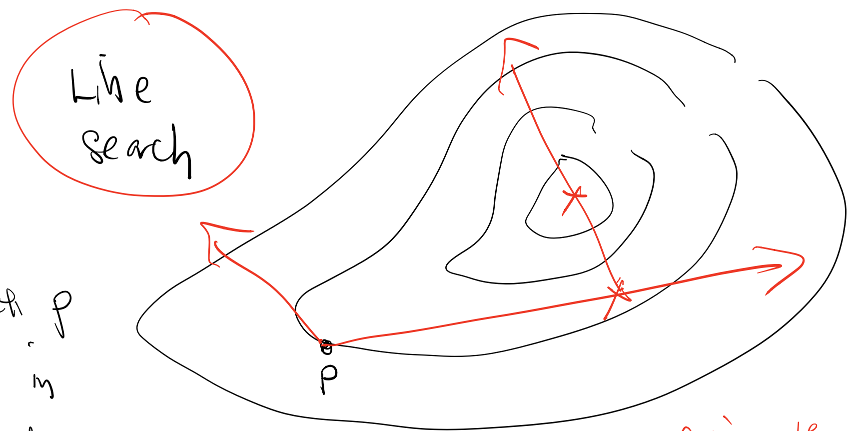
③ reflect & expand





⑤ Multiple contraction

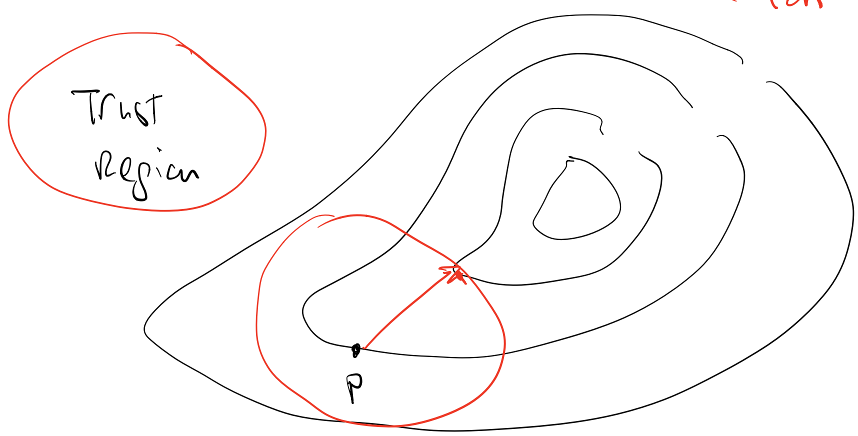
2 Approaches to continuous optimization



① Choose direction p

② Take a step in direction of p

3 interesting directions → conjugate → steepest descent → Newton



① Find a trust region
(typically this is a region where H is positive definite

Go to best location on boundary of trust region

Conjugate vectors

$\hat{p} =$
- conjugate direction
- steepest descent
- Newton direction

$$f(x) = f(p) + x^T \nabla f_p + \frac{1}{2} x^T H_p x$$

$$b = -\nabla f|_p \qquad A = H|_p$$

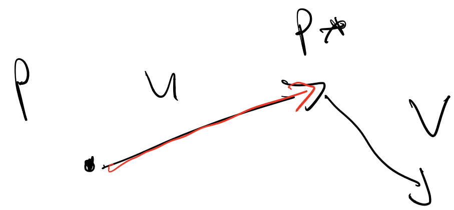
$$f(x) = c - x^T b + \frac{1}{2} x^T A x$$

$$\nabla f = -b + \frac{1}{2}(Ax + A^T x)$$

$$= Ax - b \quad (A \text{ is symmetrical})$$

Hessian

$\nabla f|_p \rightarrow$ Just a fixed vector



Fun

Gradient change / $\nabla f = Ax - b$

$$\delta(\nabla f) = A(\delta x)$$

We want the direction v to be $\perp$ to the gradient after moving along u

$\Rightarrow$ want change in gradient $\perp$ to u

$$0 = u^T \delta(\nabla f) = u^T A v$$

When

$$u^T A v = 0$$

u, v are conjugate vectors

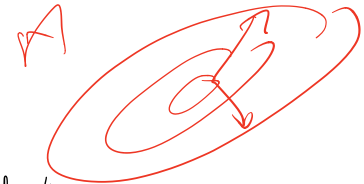
Conjugacy

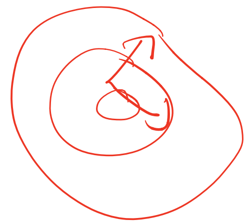
Abstractly, it is a generalization of the inner product

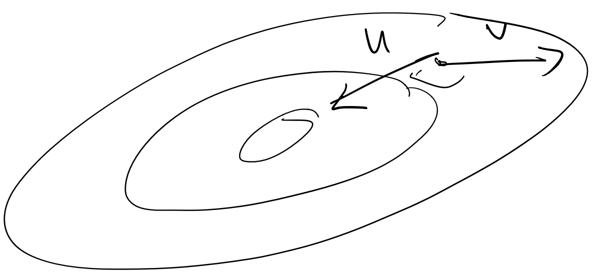
$$\langle u, v \rangle = u^T A v$$

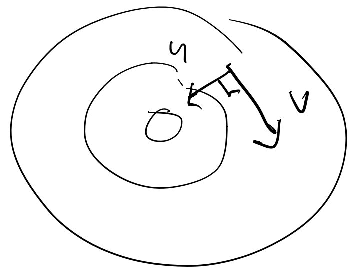
c.f. standard inner product

$$\langle u, v \rangle = u^T I v = u^T v$$









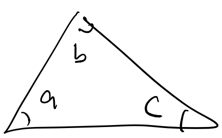
$u^T A v = 0$

In original coordinate

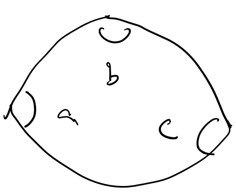
$u^T v = 0$

In standard coordinates

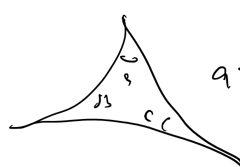
ie. $A = I$



$a+b+c = 180°$



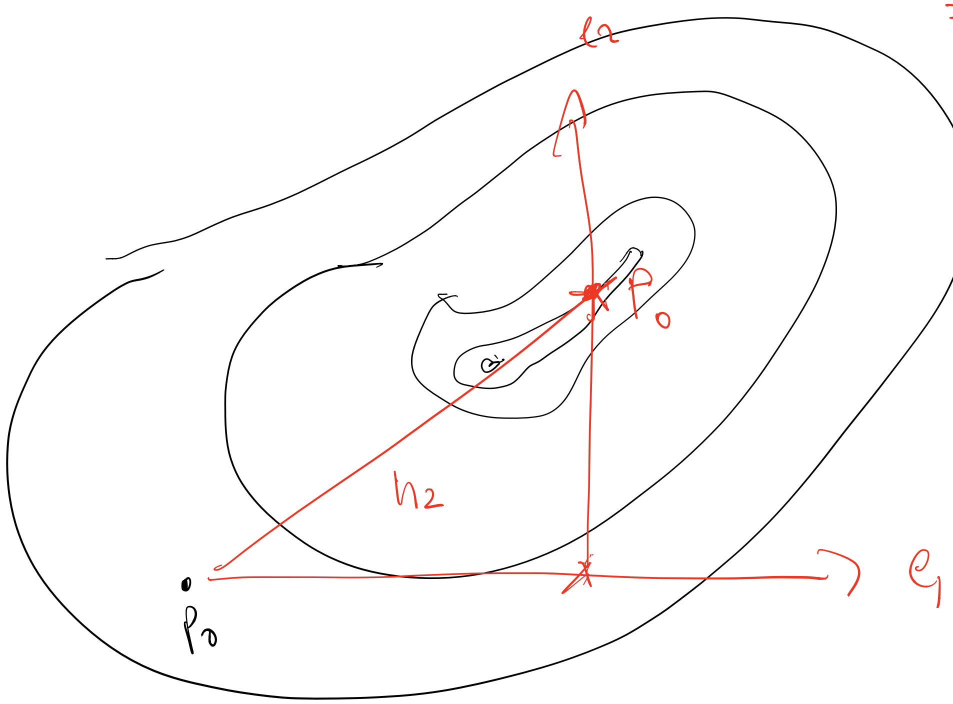
$a+b+c > 180°$



$a+b+c < 180°$

Powell's method

$\underline{N} = 2$



Powell's method finds conjugate vectors.

Newton Conjugate Gradient Algorithm

Explicit construction using Gram-Schmidt for conjugate vectors i.e given $\{u_1, u_2, u_3, \dots\}$

Find $\{v_1, v_2, v_3, \dots\}$ such that

$v_1^T v_2 = 0$

$$v_i^T A v_j = 0 \quad \forall \; i \neq j$$

$v_1^T A v_2 = 0$

(Homework Assignment)

Steepest descent

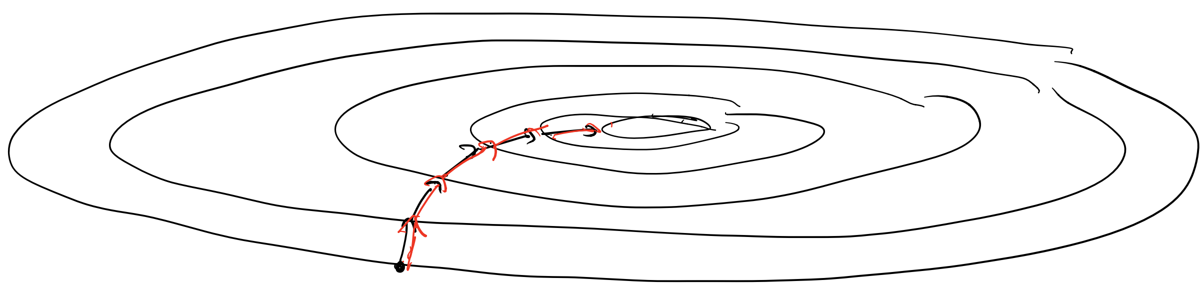
$$p_k = -\nabla f_k$$

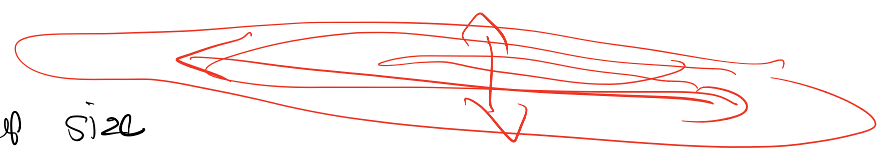
→ gradient descent

$$x_{k+1} = x_k - \alpha \nabla f_k$$

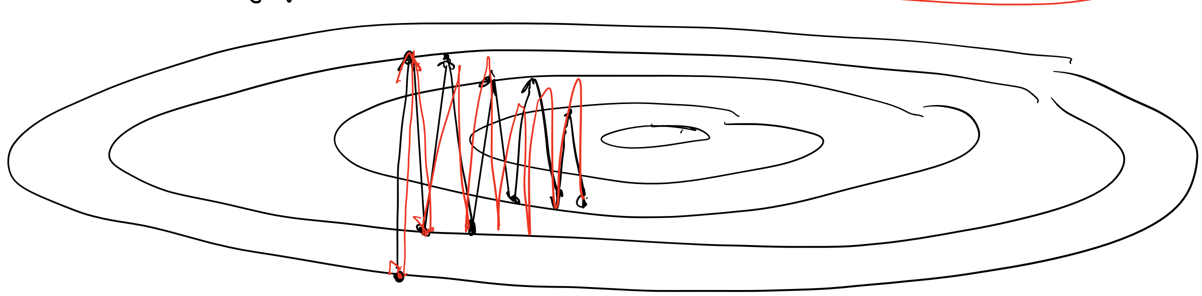
α → learning rate or step size

Good step size





Bad step size



we'll cover Newton in next lecture

C.N.

Newton direction

$$p_k = -H_k^{-1}\nabla f_k$$

Recall Newton-Raphson

$$x_{k+1} = x_k - \underbrace{\frac{f'(x_k)}{f''(x_k)}}_{p_k}$$

In higher dimensions

$$x_{k+1} = x_k \underbrace{- H_k^{-1}\nabla f_k}_{\text{Newton direction}}$$

Geometry of solution

<u>Warning</u>: Far from a minimum H_p may not be positive definite, and the Newton direction may go <u>UPHILL</u>.