Target: solve $Ax = b$

the residual function is

$$r(x) = Ax - b$$

- if r is a gradient field, then zeros of r are stationary points
- for a vector field $r(x)$ the Jacobian is A, and a gradient field has Jacobian = Hessian which implies that $r = \nabla f$ iff A is symmetric

CG setting → - furthermore if A is positive-definite then $\nabla^2 f = A > 0$ (f is strictly convex)

$$\nabla f(x) = Ax - b$$

integrate $$f(x) = \int (Ax - b) \cdot dx = \underbrace{\frac{1}{2} x^T A x - b^T x}_{\text{quadratic function}} + C$$

- in steepest descent: move in direction $-\nabla f(x)$ ⇒ zig-zagging (slow; directions are orthogonal in Euclidean sense)
- in CG: sequence of A-conjugate directions $\{p_0, p_1, \dots\}$

  A-conjugacy means $p_i^T A p_j = 0$ for $i \neq j$

  with these directions, minimization converges in at most n steps

The search begins: step size

$x_{k+1} = x_k + \alpha_k p_k$ (α_k is optimal step size)

$$f(x_{k+1}) \Rightarrow f(x_k + \alpha_k p_k) = \frac{1}{2}(x_k + \alpha_k p_k)^T A (x_k + \alpha_k p_k) - b^T (x_k + \alpha_k p_k)$$

$$\frac{df}{d\alpha_k} = p_k^T [A(x_k + \alpha_k p_k) - b]$$

Setting $\frac{df}{d\alpha_k} = 0$ gives optimal step size $\alpha_k = \frac{r_k^T p_k}{p_k^T A p_k}$ where $r_k = b - Ax_k$ (the residual)

(if $p_k = r_k$ we have steepest descent)

direction update

the new residual is:

$$r_{k+1} = b - Ax_{k+1}$$
$$r_{k+1} = b - A(x_k + \alpha_k p_k)$$
$$= \underbrace{b - Ax_k}_{r_k} - \alpha_k A p_k$$
$$r_{k+1} = r_k - \alpha_k A p_k$$

the next direction must be A-conjugate

$$p_{k+1}^T A p_k = 0$$

and also from a Krylov subspace:

$$K_k(A, r_0) = \text{span}\{r_0, Ar_0, A^2 r_0, \dots, A^{k-1} r_0\}$$

Concretely p_{k+1} must come from the newest residual and some combination of older directions:

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

$$(r_{k+1} + \beta_k p_k) A p_k = 0$$

⋮

$$\beta_k = -\frac{r_{k+1}^T A p_k}{p_k^T A p_k} \underset{\text{by orth. of residuals}}{=} \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$$