$$f(x) = \frac{1}{2}\left(Ax-b\right)^T\left(Ax-b\right) = \frac{1}{2}r^T r \tag{1}$$

$$\nabla f(x) = A^T\left(Ax-b\right) = -A^T r \tag{2}$$

$$\begin{aligned} x_{(1)} &= x_{(0)} - \alpha \nabla f(x_{(0)}) = && (3)\\ &= x_{(0)} + \alpha A^T r_{(0)} && (4) \end{aligned}$$

$$\begin{aligned} \frac{d}{d\alpha} f(x_{(1)}) &= f'(x_{(1)})^T \frac{d}{d\alpha} x_{(1)} = && (5)\\ &= f'(x_{(1)})^T A^T r_{(0)} = && (6)\\ &= (A^T r_{(1)})^T A^T r_{(0)} = && (7)\\ &= r_{(1)}^T A A^T r_{(0)} = 0 && (8)\\ & && (9) \end{aligned}$$

$$\begin{aligned} r_{(1)}^T A A^T r_{(0)} &= 0\\ \left(b - Ax_{(1)}\right)^T A A^T r_{(0)} &= 0\\ \left(b - A\left(x_{(0)} + \alpha A^T r_{(0)}\right)\right)^T A A^T r_{(0)} &= 0\\ \left(b - Ax_{(0)}\right)^T A A^T r_{(0)} - \alpha\left(A A^T r_{(0)}\right)^T A A^T r_{(0)} &= 0\\ r_{(0)}^T A A^T r_{(0)} &= \alpha\left(A A^T r_{(0)}\right)^T A A^T r_{(0)}\\ r_{(0)}^T A A^T r_{(0)} &= \alpha r_{(0)}^T A A^T A A^T r_{(0)}\\ \left(A^T r_{(0)}\right)^T\left(A^T r_{(0)}\right) &= \alpha\left(A^T r_{(0)}\right)^T A^T A\left(A^T r_{(0)}\right)\\ \alpha &= \frac{\left(A^T r_{(0)}\right)^T\left(A^T r_{(0)}\right)}{\left(A^T r_{(0)}\right)^T A^T A\left(A^T r_{(0)}\right)} \end{aligned}$$