1. Convergence and the Error
   a. The error looks like $e_0 = \sum_j a_j s_j$ where the $s_j$ are the search directions and $a_j$ are numerical coefficients
   b. $s_k \cdot A e_0 = s_k \cdot A \sum_j a_j s_j = \sum_j a_j s_k \cdot A s_j = a_k s_k \cdot A s_k$ since the search directions are orthogonal in A space
      i. thus $a_k = \frac{s_k \cdot A e_0}{s_k \cdot A s_k} = \frac{s_k \cdot A (e_0 + \sum_{j=1}^{k-1} \alpha_j s_j)}{s_k \cdot A s_k}$ where the summation can be added since it is identically zero when multiplied by $s_k \cdot A$
      ii. now $e_k = e_{k-1} + s_{k-1} \alpha_{k-1} = e_{k-2} + s_{k-2} \alpha_{k-2} + s_{k-1} \alpha_{k-1} = \ldots$
         i. e. $e_k = e_0 + \sum_{j=1}^{k-1} \alpha_j s_j$
      iii. thus $a_k = \frac{s_k \cdot A e_k}{s_k \cdot A s_k}$ and (from above) $a_k = -\alpha_k$
   c. so the error is $e_0 = -\sum_j \alpha_j s_j$ and $e_k = -\sum_j \alpha_j s_j + \sum_{j=1}^{k-1} \alpha_j s_j$
      i. after n steps the second term is equal to the first term and the error is zero
   d. multiplying this error equation by $s_i \cdot A$ gives
      $s_i \cdot A e_k = -\sum_j \alpha_j s_i \cdot A s_j + \sum_{j=1}^{k-1} \alpha_j \sum j s_i \cdot A s_j$
      i. for $i < k$, there is exactly one nonzero term in each sum, and these terms cancel
      ii. thus for $i < k$, $s_i \cdot A e_k = 0$ and thus $s_i \cdot r_k = 0$ (we’ll use this below)
      iii. this means that the current residual at step $k$ is orthogonal to all the previous search directions
2. finding the A-orthogonal directions with Gram-Schmidt
   a. given a vector $V$, construct $s_k$ by subtracting out the “A-overlap” of $V$ with $s_i$ to $s_{k-1}$ so that $s_k \cdot A s_i = 0$ for $i=1,k-1$
   b. we define $s_k = V - \sum_{j=1}^{k-1} \frac{V \cdot A s_j}{s_j \cdot A s_j} s_j$
      i. note that $s_k \cdot A s_i = V \cdot A s_i - \sum_{j=1}^{k-1} \frac{V \cdot A s_j}{s_j \cdot A s_j} s_j \cdot A s_i$ and then all the terms in the sum vanish except for one leaving$ s_k \cdot A s_i = V \cdot A s_i - \frac{V \cdot A s_k}{s_k \cdot A s_i} s_k \cdot A s_i = 0$ as desired
c. for \( i \geq k \), \( s_k \cdot r_i = V_k \cdot r_i - \sum_{j=1}^{k-1} \frac{V_k \cdot A s_j}{s_j \cdot A s_j} s_j \cdot r_i = V_k \cdot r_i \), where the summation vanishes because the residual at step \( i \) is orthogonal to all the previous search directions

i. when \( k=i \) this leads to \( s_k \cdot r_k = V_k \cdot r_k \) and \( \alpha_k = \frac{s_k \cdot r_k}{s_k \cdot A s_k} = \frac{V_k \cdot r_k}{s_k \cdot A s_k} \) (we’ll use this below)

ii. when \( k < i \), \( 0 = V_k \cdot r_i \), i.e. the residual is orthogonal to all the previous \( V_k \) as well (we’ll use this below)