CS205 – Class 11

Covered in class: Everything
Readings: Shewchuk Paper on course web page and Heath 473-478

1. Conjugate Directions
   a. The goal is to choose a sequence of search directions s_0, s_1, … that are all orthogonal. Then only one step is needed in each search direction
   b. We update $x_{k+1} = x_k + s_k \alpha_k$ where each α_k is chosen so that all the remaining error is orthogonal to the search direction s_k. That is, $e_{k+1} \cdot s_k = 0$ and we never need to step in the s_k direction again.
      i. Note that $x_{k+1} = x_k + s_k \alpha_k$ leads to $x_{k+1} - x_{exact} = x_k - x_{exact} + s_k \alpha_k$ or $e_{k+1} = e_k + s_k \alpha_k$ (we'll use this later too!)
      ii. $e_{k+1} \cdot s_k = 0$ leads to $\left(e_k + s_k \alpha_k\right) \cdot s_k = 0$ or $\alpha_k = -\dfrac{e_k \cdot s_k}{s_k \cdot s_k}$
         1. however we don't know e_k
   c. Just as before, instead of using e, we can use r=-Ae to obtain $r_{k+1} \cdot s_k = 0$ or $Ae_{k+1} \cdot s_k = 0$
      i. This leads to $A\left(e_k + s_k \alpha_k\right) \cdot s_k = 0$ or $\alpha_k = -\dfrac{s_k \cdot Ae_k}{s_k \cdot As_k} = \dfrac{s_k \cdot r_k}{s_k \cdot As_k}$
      ii. When $s_k = r_k$ this is the steepest decent method
   d. The only issue now is how to choose the s_k, and the trick of conjugate directions is to choose them *A-orthogonal* instead of orthogonal
      i. That is, $s_j \cdot As_k = 0$ for $j \neq k$ (instead of $s_j \cdot s_k = 0$)
      ii. If A is symmetric, $0^T = \left(s_j \cdot As_k\right)^T = \left(s_j^T As_k\right)^T = s_k^T As_j = s_k \cdot As_j$ and the A-orthogonal relationship is symmetric

2. Why does A-orthogonal work?
   a. The error looks like $e_0 = \sum_j a_j s_j$ where the s_j are the search directions that span the space and a_j are numerical coefficients
   b. $s_k \cdot Ae_0 = s_k \cdot A \sum_j a_j s_j = \sum_j a_j s_k \cdot As_j = a_k s_k \cdot As_k$ since the search directions are orthogonal in A space
      i. thus $a_k = \dfrac{s_k \cdot Ae_0}{s_k \cdot As_k} = \dfrac{s_k \cdot A(e_0 + \sum_{j=1}^{k-1} \alpha_j s_j)}{s_k \cdot As_k}$ where the summation can be added since it is identically zero when multiplied by $s_k \cdot A$
      ii. now $e_k = e_{k-1} + s_{k-1} \alpha_{k-1} = e_{k-2} + s_{k-2} \alpha_{k-2} + s_{k-1} \alpha_{k-1} = \ldots$
         1. i.e. $e_k = e_0 + \sum_{j=1}^{k-1} \alpha_j s_j$
      iii. thus $a_k = \dfrac{s_k \cdot Ae_k}{s_k \cdot As_k}$ 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

3. Aside: Multiplying this error equation by $s_i \cdot A$ gives $s_i \cdot Ae_k = -\sum_j \alpha_j s_i \cdot As_j + \sum_{j=1}^{k-1} \alpha_j s_i \cdot As_j$

a. for i < k, there is exactly one nonzero term in each sum, and these terms cancel

b. thus for i < k, $s_i \cdot Ae_k = 0$ and thus $s_i \cdot r_k = 0$ (we'll use this below)

c. this means that the current residual at step k is orthogonal to all the previous search directions