Exercise 6

Lorenzo Vigo

Optimization: 1^{st} November 2021

Exercise 6. To be delivered before 2-XI-2021 as: `Ex06-YourSurname.pdf`

Consider the conjugate gradient method applied to the minimization of

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

where A is a positive definite and symmetric matrix.
Show that the iterate $\boldsymbol{x}^k$ minimizes f over

$$\boldsymbol{x}^0 + < \boldsymbol{v}^0, A\boldsymbol{v}^0, ..., A^{k-1}\boldsymbol{v}^0 >$$

where $\boldsymbol{v}^0 = \nabla f(\boldsymbol{x}^0)$, and $< v^0, A\boldsymbol{v}^0, ..., A^{k-1}\boldsymbol{v}^0 >$ is the subspace generated by $v^0, A\boldsymbol{v}^0, ..., A^{k-1}\boldsymbol{v}^0$

Following the Conjugate Gradient Method Theorem seen in class, we know that the method will start at x^0 and move along $z^1, ..., z^k$ such that $(z^j)^T \nabla f(x^j) = (z^j)^T \nabla f(x^k) = 0, j = 1, ..., k$. This means that the method minimizes f over $x^0 + < z^1, ..., z^k >$.

All we need to prove is that $< z^1, ..., z^k > \subset < v^0, Av^0, ... A^{k-1}v^0 >$. Following the choice of the conjugate directions shown in class, we know that:

$$z^1 = -\nabla f(x^0) = -v^0$$

Therefore $z^1 \in < v^0, Av^0, ..., A^{k-1}v^0 >$. Let's suppose we have already generated $z^1, ..., z^m$ for $1 < m < k$ such that $z^i \in < v^0, Av^0, ..., A^{k-1}v^0 >$ for $i = 1, ..., m$. We know that, for a specific β:

$$z^{m+1} = -\nabla f(x^m) + \beta z^m$$

As we are supposing that $z^i \in < v^0, Av^0, ..., A^{k-1}v^0 >$ for $i = 1, ..., m$, we just need to prove that $\nabla f(x^m)$ is included in the same generated space.

We should note that when A is positive definite and symmetric:

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

Therefore:

$$
\begin{aligned}
\nabla f(x^m) &= Ax^m - b \\
\nabla f(x^m) &= A(x^{m-1} + \alpha_m z^m) - b \\
\nabla f(x^m) &= (Ax^{m-1} - b) + \alpha_m A z^m \\
&\ldots \\
\nabla f(x^m) &= (Ax^0 - b) + \sum_{j=1}^{m} \alpha_j A z^j \\
\nabla f(x^m) &= \nabla f(x^0) + \sum_{j=1}^{m} \alpha_j A z^j \\
\nabla f(x^m) &= v^0 + \sum_{j=1}^{m} \alpha_j A z^j
\end{aligned}
\tag{1}
$$

We can now state that $f(x^m) \in < v^0, Av^0, ..., A^{k-1}v^0 >$, which proves the property given in the exercise.