These equations imply that

$$\|e_k\|_A \le \left(\inf_{P\in\mathcal{P}_k} \max_{1\le i\le n} |P(\lambda_i)|\right) \|e_0\|_A .$$

It can be shown that the conjugate gradient method requires of the order of

n^3 additions,

n^3 multiplications,

$2n$ divisions.

In theory, this is worse than the number of elementary operations required by the Cholesky method. Even though the conjugate gradient method does not seem to be the best method for *full* matrices, it usually outperforms other methods for *sparse* matrices. The reason is that the matrix A only appears in the computation of the vector Ad_k. If the matrix A is banded (for example, tridiagonal), computing Ad_k is very cheap and there is no need to store the entire matrix A, in which case the conjugate gradient method is fast. Also, although in theory, up to n iterations may be required, in practice, convergence may occur after a much smaller number of iterations.

Using the inequality

$$\|e_k\|_A \le \left(\inf_{P\in\mathcal{P}_k} \max_{1\le i\le n} |P(\lambda_i)|\right) \|e_0\|_A ,$$

by choosing P to be a shifted Chebyshev polynomial, it can be shown that

$$\|e_k\|_A \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|e_0\|_A ,$$

where $\kappa = \mathrm{cond}_2(A)$; see Trefethen and Bau [176] (Lecture 38, Theorem 38.5). Thus the rate of convergence of the conjugate gradient method is governed by the ratio

$$\frac{\sqrt{\mathrm{cond}_2(A)}-1}{\sqrt{\mathrm{cond}_2(A)}+1},$$

where $\mathrm{cond}_2(A) = \lambda_n/\lambda_1$ is the condition number of the matrix A. Since A is positive definite, λ_1 is its smallest eigenvalue and λ_n is its largest eigenvalue.

The above fact leads to the process of *preconditioning*, a method which consists in replacing the matrix of a linear system $Ax = b$ by an "equivalent" one for example $M^{-1}A$ (since M is invertible, the system $Ax = b$ is equivalent to the system $M^{-1}Ax = M^{-1}b$), where M is chosen so that $M^{-1}A$ is still symmetric positive definite and has a smaller condition number than A; see Trefethen and Bau [176] (Lecture 40) and Demmel [48] (Section 6.6.5).