If $A = A^T$ is positive definite, there is an even more efficient Krylov subspace method than MINRES, namely the method of *Conjugate Gradients* (for $Ax = b$) which computes $x_k \in x_0 + \mathcal{K}_k(A, r_0)$ such that $\|x - x_k\|_A$ is minimal where $y^T A y = \|y\|_A^2$.

Note $\|\cdot\|_A$ defines a norm when A is symmetric and positive definite, indeed $< z, y >_A = < Az, y > = y^T A z$ defines a scalar product (inner product) for which $< y, y >_A = \|y\|_A^2$ in this case.

Lemma Conjugate Gradient Algorithm

choose x_0 , $r_0 = b - Ax_0 = p_0$ and for $k = 0, 1, 2, \ldots$

$$
\begin{aligned}
\alpha_k &= p_k^T r_k / p_k^T A p_k \\
x_{k+1} &= x_k + \alpha_k p_k \\
r_{k+1} &= b - A x_{k+1} \\
\beta_k &= -p_k^T A r_{k+1} / p_k^T A p_k \\
p_{k+1} &= r_{k+1} + \beta_k p_k
\end{aligned}
$$

computes iterates $\{x_k\}$, corresponding residuals $\{r_k\}$ and search directions $\{p_k\}$ so that as long as $x_k \neq x$ we have

$$
\begin{aligned}
r_k^T p_j &= r_k^T r_j = 0 \ , \ j < k \qquad (1) \\
p_k^T A p_j &= 0 \ , \ j < k \ , \ (p_k^T A p_k \neq 0) \ (2) \\
\text{span}\{r_0, r_1, \ldots, r_{k-1}\} &= \text{span}\{p_0, p_1, \ldots, p_{k-1}\} \\
&= \mathcal{K}_k(A, r_0) \qquad (3)
\end{aligned}
$$

It is now an easy induction using $x_k = x_{k-1} + \alpha_{k-1} p_{k-1}$ to show that $x_k \in x_0 + \mathcal{K}_k(A, r_0)$ (exercise)
and also

Theorem

$$\|x - x_k\|_A \leq \|x - y\|_A \quad , \quad y \in x_0 + \mathcal{K}_k(A, r_0)$$

Proof
let $c = x - x_0$ and $c_k = x_k - x_0 \in \mathcal{K}_k(A, r_0)$
then $Ac = r_0$ and $x - x_k = c - c_k$
hence $r_k = A(x - x_k) = A(c - c_k)$.

Now by the above r_k is orthogonal to every vector in $\mathcal{K}_k(A, r_0)$ i.e. $\forall\, v \in \mathcal{K}_k(A, r_0)$

$$0 = \langle r_k, v \rangle = \langle A(c - c_k), v \rangle = \langle c - c_k, v \rangle_A$$

and such (Galerkin) orthogonality $\Rightarrow \|c - y\|_A$ is minimised for $y \in \mathcal{K}_k(A, r_0)$ when $y = c_k$

$\Rightarrow \|x - z\|_A$ is minimised for $z \in x_0 + \mathcal{K}_k(A, r_0)$ by $z = x_k$ since $c = x - x_0$ and $x_k = x_0 + c_k$.

Convergence of Conjugate Gradients

we have $r_k = p_k(A)r_0$ or equivalently $(x - x_k) = p_k(A)(x - x_0)$, $p_k \in \Pi_k, p_k(0) = 1$ and $\|x - x_k\|_A = \|r_k\|_{A^{-1}}$ minimal over $x_k \in x_0 + \mathcal{K}_k(A, r_0)$ for each k.

Let $Av_j = \lambda_j v_j, \quad j = 1, \ldots, n, \quad v_j^T v_i = \delta_{ij}$ and $x - x_0 = \sum_{j=1}^n \alpha_j v_j$ then

$$x - x_k = \sum_{j=1}^n \alpha_j p_k(A) v_j = \sum_{j=1}^n \alpha_j p_k(\lambda_j) v_j$$

so $$\langle x - x_k, x - x_k \rangle_A = \sum_{j=1}^n \alpha_j^2 p_k(\lambda_j)^2 \langle v_j, v_j \rangle_A$$

since $v_j^T v_i = \delta_{ij} \Rightarrow v_j^T A v_i = \lambda_i \delta_{ij} = 0$ if $i \neq j$

hence $$\|x - x_k\|_A \leq \min_{p \in \Pi_k, p(0)=1} \max_j |p(\lambda_j)| \, \|x - x_0\|_A$$