Dado que para $v_1, v_2 \in \mathbb{R}^n$, $\langle v_1, v_2 \rangle_A = 0$ significa que son ortogonales, podemos crear una base a partir de estos vectores. Es decir,

$\beta = \{p_1, \dots, p_n\}$ forman una base para $\langle p_i, p_j \rangle_A = 0 \quad \forall i,j \in \{1, \dots, n\}$

Luego, la solución x al sistema $Ax = b$ puede escribirse cambiando de base como $\tilde{x} = \alpha_1 p_1 + \dots + \alpha_n p_n$, con $\alpha_i \in \mathbb{R} \ \forall i \in \{1, \dots, n\}$.

De esto se sigue que $b = A\tilde{x} = A(\alpha_1 p_1 + \dots + \alpha p_n) = \alpha_1 A p_1 + \dots + \alpha_n A p_n$

Luego, para $i \in \{1, \dots, n\}$, $p_i^T b = p_i^T(\alpha_1 A p_1 + \dots + \alpha_n A p_n)$

$= p_i^T \alpha_1 A p_1 + \dots + p_i^T \alpha_i A p_i + \dots + p_i^T \alpha_n A p_n = \alpha_1 p_i^T A p_1 + \dots + \alpha_i p_i^T A p_i + \dots$

$+ \alpha_n p_i^T A p_n = \alpha_1 \langle p_i, p_1 \rangle_A + \dots + \alpha_i \langle p_i, p_i \rangle_A + \dots + \alpha_n \langle p_i, p_n \rangle_A$

Luego, por ortogonalidad de los vectores de la base, tenemos que $p_i^T b = 0 + \dots + \alpha_i \langle p_i, p_i \rangle_A + 0 + \dots + 0 = \alpha_i \langle p_i, p_i \rangle_A$.

Así, $\alpha_i = \dfrac{p_i^T b}{\langle p_i, p_i \rangle_A} = \dfrac{p_i^T b}{p_i^T A p_i}$

Ahora, observe que para el método iterativo se parte de un vector x_1 y se define el residuo $r_1 = Ax_1 - b$. Podemos interpretar esto como $p_1 = b - Ax_1 = -r_1$. Luego, para una iteración k el residuo se define como $r_k = b - Ax_k$

Podemos continuar por el método de ortogonalización de Gram-Schmidt para extender la base β. Así,

$$p_k = r_k - \frac{p_k^T A r_k}{\underbrace{p_k^T A p_k}_{= \beta_k}} p_k$$ (ortonormalizando). Luego,

$x_{k+1} = x_k + \alpha_k p_k$ / dado que se busca llegar a la forma $\tilde{x}$.

Así, retomando $*$, vemos que:

$$p_k^T = (r_k + \beta_{k-1} p_{k-1})^T$$

$$Ap_k = \frac{1}{\alpha_k}(r_k - r_{k+1})$$

luego $p_k^T A p_k = (r_k + \beta_{k-1} p_{k-1})^T A p_k = \frac{1}{\alpha_k} r_k^T r_k$

$$\Rightarrow \qquad \alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}$$