Ex 3.25

The correlation between the active set A_K and the residual is given by $X_{A_K}^T r_K(\alpha)$. Previously we have incrementally increased α until some variable not in the active set reaches the same correlation with the residual. In this question we will attempt to find this new variable and its associated α algebraically.

We showed in Ex 3.23 (a) that the correlations between the active set and the residuals are tied, even as we change α. Thus we need only calculate $x_a^T r_K(\alpha)$ where x_a is any member of the active set. Additionally we denote x_b for a variable that is not a member of the active set.

Now, we assume step K is complete and we are interested in the variable added and α value in the next step K+1. Therefore we will examine:

$$
\begin{aligned}
& x_a^T r_{K+1}(\alpha) \\
= \; & x_a^T(y - X_{A_K}\beta_{A_K}(\alpha)) && \text{(P. 74)} \\
= \; & x_a^T(y - X_{A_K}\beta_{A_K} - \alpha X_{A_K}\delta_K) && \text{(P. 74)} \\
= \; & x_a^T(r_K - \alpha X_{A_K}\delta_K) && \text{(P. 74)} \\
= \; & x_a^T r_K - \alpha x_a^T X_{A_K}\delta_K
\end{aligned}
$$

and identically we find that:

$$
\begin{aligned}
& x_b^T r_{K+1}(\alpha) \\
= \; & x_b^T r_K - \alpha x_b^T X_{A_K}\delta_K
\end{aligned}
$$

Now when $\alpha = \alpha^*$, the stopping value, we know that the variable to add to the active set is the one with the most correlation with the residual:

$$\Rightarrow \max_b \left(\left| x_b^T r_{K+1}(\alpha^*) \right| \right)$$

but since we do not yet know the stopping value we cannot solve this. We can, however, find what value α would take for each x_b by solving:

$$x_a^T r_{K+1}(\alpha) = \left| x_b^T r_{K+1}(\alpha) \right|$$
$$= x_a^T r_K - \alpha x_a^T X_{A_K} \delta_K = \left| x_b^T r_K - \alpha x_b^T X_{A_K} \delta_K \right|$$

This splits into two cases depending on the sign of the RHS:

① Positive case: $x_a^T r_K - \alpha x_a^T X_{A_K} \delta_K = x_b^T r_K - \alpha x_b^T X_{A_K} \delta_K$

$$\Rightarrow x_a^T r_K - x_b^T r_K = \alpha \left(-x_b^T X_{A_K} \delta_K + x_a^T X_{A_K} \delta_K \right)$$

$$\Rightarrow \alpha = \frac{x_a^T r_K - x_b^T r_K}{-x_b^T X_{A_K} \delta_K + x_a^T X_{A_K} \delta_K}$$

② Negative case: $x_a^T r_K - \alpha x_a^T X_{A_K} \delta_K = -x_b^T r_K + \alpha x_b^T X_{A_K} \delta_K$

$$\Rightarrow x_a^T r_K + x_b^T r_K = \alpha \left(x_b^T X_{A_K} \delta_K + x_a^T X_{A_K} \delta_K \right)$$

$$\Rightarrow \alpha = \frac{x_a^T r_K + x_b^T r_K}{x_b^T X_{A_K} \delta_K + x_a^T X_{A_K} \delta_K}$$

Now since α increases from 0 to 1 the next variable to add to the active set is the minimum α across all x_b. Thus the above equations find the next variable and its α value at any step K.