4. \( \Sigma(AX_iBY_j)_{ij} = \text{Cov}(AX_i, BY_j) = \text{Cov}(\sum_u A_{iu} X_u, \sum_v B_{jv} Y_v) \)
\[
= \sum_u \sum_v A_{iu} B_{jv} \text{Cov}(X_u, Y_v) = \sum_u \sum_v A_{iu} B_{jv} \Sigma(X_u, Y_v)
\]
\[
= (A \Sigma(X,Y) B^t)_{ij}
\]

(b) \( \text{Var}(AX) = \Sigma(AX, AX) = A \Sigma(X, X) A^t = A \text{Var}(X) A^t \)

(c) \( \text{Var}(AX + BY) \) only makes sense if

\[A \text{ is } m \times n\]
\[X \text{ is } n \times 1\]
\[B \text{ is } m \times k\]
\[Y \text{ is } k \times 1\]

for some choice of \( m, n, k \). In this case we have

\[\text{Var}(AX + BY) = A \text{Var}(X) A^t + B \text{Var}(Y) B^t + 2 A \Sigma(X,Y) B^t\]

To see this, we'll need a couple of simple properties of \( \Sigma \):

1. \( \Sigma(AX_iY_j) = A \Sigma(X_i, Y_j) \)
2. \( \Sigma(X_i BY_j) = \Sigma(Y_j, X_i) B^t \)
3. \( \Sigma(X_i Y_i) = \Sigma(Y_i, X_i)^t \)
4. \( \Sigma(X_i + X_j Y_i) = \Sigma(X_i, Y_i) + \Sigma(X_j, Y_i) \)
5. \( \Sigma(X_i Y_i + Y_j) = \Sigma(X_i Y_i) + \Sigma(Y_j, Y_j) \)

Then we have

\[\text{Var}(AX + BY) = \Sigma(AX + BY, AX + BY)\]
\[
= A \Sigma(X, X) A^t + B \Sigma(Y, Y) B^t + A \Sigma(X, Y) B^t
\]
\[
+ B \Sigma(Y, X) A^t
\]
\[
= A \text{Var}(X) A^t + B \text{Var}(Y) B^t + A \Sigma(X, Y) B^t + (A \Sigma(X, Y) B^t)^t
\]