A7

(a)

We expand $f(x, y)$ as:

$$f(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_i x_j + \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} y_i x_j + c.$$

(b)

The gradient of $f(x, y)$ with respect to x is given by:

$$\nabla_x f(x, y) = (A + A^\top)\, x \;+\; B^\top y.$$

In summation form, the k-th component of the gradient is:

$$\frac{\partial f}{\partial x_k} = \sum_{j=1}^{n} (A_{kj} + A_{jk})\, x_j + \sum_{i=1}^{n} B_{ik}\, y_i.$$

This arises because A is not necessarily symmetric; thus the partial derivatives from both $A_{kj} x_k x_j$ and $A_{ik} x_i x_k$ add up to $\sum_j (A_{kj} + A_{jk}) x_j$, rather than $2 \sum_j A_{kj} x_j$.

This is derived as follows:

1. **For the term** $x^\top A x = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_i x_j$:

$$\text{When we take } \frac{\partial}{\partial x_k}, \text{only terms where } i = k \text{ or } j = k \text{ contribute.}$$

$$\text{If } i = k,\; A_{kj} x_k x_j \quad \Longrightarrow \quad \frac{\partial}{\partial x_k}(A_{kj} x_k x_j) = A_{kj} x_j.$$

$$\text{If } j = k,\; A_{ik} x_i x_k \quad \Longrightarrow \quad \frac{\partial}{\partial x_k}(A_{ik} x_i x_k) = A_{ik} x_i.$$

   Adding these contributions:

$$\sum_{j=1}^{n} A_{kj} x_j \;+\; \sum_{i=1}^{n} A_{ik} x_i \;=\; \sum_{j=1}^{n} (A_{kj} + A_{jk})\, x_j,$$

   i.e. $\left[(A + A^\top)\, x\right]_k$.

2. **For the term** $y^\top B x = \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij} y_i x_j$:

$$\text{When we take } \frac{\partial}{\partial x_k}, \text{only terms where } j = k \text{ contribute.}$$

$$\text{If } j = k,\; B_{ik} y_i x_k \quad \Longrightarrow \quad \frac{\partial}{\partial x_k}(B_{ik} y_i x_k) = B_{ik} y_i.$$

   Summing over i yields: $\sum_{i=1}^{n} B_{ik} y_i$, i.e. $\left[B^\top y\right]_k$.

3. **For the constant term** c:

$$\frac{\partial}{\partial x_k}\, c = 0.$$

Combining these results for each component k:

$$\frac{\partial f}{\partial x_k} = \left[(A + A^\top)\, x\right]_k \;+\; \left[B^\top y\right]_k.$$

Hence in vector form:

$$\nabla_x f(x, y) = (A + A^\top)\, x + B^\top y.$$

(c)

The gradient of $f(x, y)$ with respect to y is:

$$\nabla_y f(x, y) = B\,x.$$

In summation form, the k-th component of the gradient is:

$$\frac{\partial f}{\partial y_k} = \sum_{j=1}^{n} B_{kj} x_j.$$

This follows because $x^\top Ax$ does not involve y. Taking partials of $y^\top Bx$ with respect to y_k isolates exactly those terms $B_{kj}x_j$.