Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be defined by
\[
f(x) = \frac{x^T A x}{x^T x},
\]
where \( A \) is a symmetric matrix, and \( x \in \mathbb{R}^n \) is represented by a column vector. This function is called a Rayleigh quotient.

- Rewrite the formula for \( f(x) \) in terms of dot products.

**Solution**
\[
f(x) = \frac{x \cdot A x}{x \cdot x}
\]

- Let \( x = (x_1, x_2, \ldots, x_n) \) and let \( A \) have entries \( a_{ij}, \) for \( i, j = 1, 2, \ldots, n. \) Rewrite the formula for \( f(x) \) in terms of the components of \( x \) and the entries of \( A. \)

**Solution** Using the definition of matrix-vector multiplication, we obtain
\[
f(x) = \frac{\sum_{i=1}^{n} x_i [Ax]_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} x_j}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j}{\sum_{i=1}^{n} x_i^2}.
\]

- Compute \( \nabla f. \) *Hint:* Use the fact that differentiation rules from single-variable calculus generalize to the gradient. See the Lecture 7 Notes.

**Solution** From the Quotient Rule,
\[
\nabla f = \frac{(x^T x) \nabla(x^T A x) - (x^T A x) \nabla(x^T x)}{(x^T x)^2}.
\]

It follows that for each \( k = 1, 2, \ldots, n, \) we have
\[
\frac{\partial f}{\partial x_k} = \frac{\left( \sum_{i=1}^{n} x_i^2 \right) \left( \sum_{j=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \right) x_k - \left( \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \right) \left( \sum_{i=1}^{n} x_i^2 \right) x_k}{\left( \sum_{i=1}^{n} x_i^2 \right)^2}.
\]

\[
= \frac{\left( \sum_{i=1}^{n} x_i^2 \right) \left( \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i \right) - \left( \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \right) (2x_k)}{\left( \sum_{i=1}^{n} x_i^2 \right)^2}.
\]

\[
= \frac{\left( \sum_{i=1}^{n} x_i^2 \right) \left( \sum_{j=1}^{n} a_{kj} x_j + \sum_{j=1}^{n} a_{jk} x_j \right) - \left( \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \right) (2x_k)}{\left( \sum_{i=1}^{n} x_i^2 \right)^2}.
\]
\[
\begin{align*}
&= \left( \sum_{i=1}^n x_i^2 \right) \left( 2 \sum_{j=1}^n a_{kj} x_j \right) - \left( \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \right) (2x_k) \\
&= \frac{2 \sum_{i=1}^n x_i^2}{\left( \sum_{i=1}^n x_i^2 \right)^2} \\
&= \frac{2}{x^T x} \frac{x^T[Ax]_k - (x^T Ax) x_k}{(x^T x)^2} \\
&= \frac{2}{x^T x} \left( [Ax]_k - \frac{x^T Ax}{x^T x} x_k \right).
\end{align*}
\]

It follows that
\[
\nabla f(x) = \frac{2}{x^T x} \left( Ax - \frac{x^T Ax}{x^T x} x \right) = \frac{2}{x^T x} (Ax - f(x)x).
\]

- How do the critical points of \( f \) relate to the eigenvalues of \( A \)?

**Solution** In order for \( \nabla f(x) = 0 \), we must have
\[
Ax = \lambda x,
\]
where
\[
\lambda = f(x) = \frac{x^T Ax}{x^T x}.
\]
The relation \( Ax = \lambda x \) occurs if \( x \) is an eigenvector of \( A \), with corresponding eigenvalue \( \lambda \). But then we have
\[
f(x) = \frac{x^T Ax}{x^T x} = \frac{x^T (\lambda x)}{x^T x} = \frac{\lambda x^T x}{x^T x} = \lambda,
\]
so we conclude that each eigenvector of \( A \) is a local maximum, local minimum or saddle point of \( f \), and the corresponding function values are the eigenvalues.