

# The Gradient of a Quadratic Form: An Operator Perspective

Your Friendly Math Teacher

April 6, 2025

## Abstract

This note presents a coordinate-free derivation of the gradient of a general quadratic form  $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle$  defined on a real inner product space. We emphasize the roles of the directional derivative, the inner product structure, and the adjoint operator.

## 1 Introduction: Beyond Matrices

Quadratic forms appear everywhere – from conic sections and optimization problems in calculus to energy functionals in physics and variance calculations in statistics. Often, we encounter them in the matrix form  $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ . Finding the gradient,  $\frac{\partial f}{\partial \mathbf{x}}$ , is a standard calculation, resulting in  $(A + A^T)\mathbf{x}$ , or  $2A\mathbf{x}$  if  $A$  is symmetric.

While the matrix approach is effective, especially for computation, viewing the problem through the lens of linear operators on abstract vector spaces offers deeper insights. It strips away the dependence on a specific basis and highlights the fundamental geometric and algebraic structures involved. Our goal today is to derive the gradient of  $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle$  using this more general framework.

## 2 Setting the Stage

Let's clearly define our playground:

**Definition 2.1** (Inner Product Space). We work within a finite-dimensional real vector space  $\mathcal{V}$ , equipped with an inner product  $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ . Recall that an inner product is a symmetric, bilinear, positive-definite form. It provides notions of length, angle, and orthogonality.

**Definition 2.2** (Linear Operator). Let  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{V}$  be a linear operator. Linearity means  $\mathbf{T}(a\mathbf{u} + b\mathbf{v}) = a\mathbf{T}\mathbf{u} + b\mathbf{T}\mathbf{v}$  for all  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$  and  $a, b \in \mathbb{R}$ .

**Definition 2.3** (Quadratic Form). The function we wish to differentiate is the quadratic form  $f : \mathcal{V} \rightarrow \mathbb{R}$  associated with  $\mathbf{T}$ , defined by:

$$f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle$$

## 3 The Gradient and the Directional Derivative

How do we define the "gradient" in this abstract setting? We use the concept of the directional derivative.

**Definition 3.1** (Directional Derivative). The directional derivative of  $f$  at  $\mathbf{x}$  in the direction  $\mathbf{h} \in \mathcal{V}$  measures the instantaneous rate of change of  $f$  as we move from  $\mathbf{x}$  along  $\mathbf{h}$ . It's defined as:

$$D_{\mathbf{h}}f(\mathbf{x}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x} + t\mathbf{h}) - f(\mathbf{x})}{t}$$

provided the limit exists.

**Definition 3.2** (Gradient). The gradient of  $f$  at  $\mathbf{x}$ , denoted  $\nabla f(\mathbf{x})$ , is the *unique* vector in  $\mathcal{V}$  such that the directional derivative can always be expressed as an inner product with  $\nabla f(\mathbf{x})$ :

$$D_{\mathbf{h}}f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{h} \rangle \quad \text{for all } \mathbf{h} \in \mathcal{V}$$

*Remark 3.1.* The existence and uniqueness of such a vector  $\nabla f(\mathbf{x})$  for differentiable functions on finite-dimensional inner product spaces is guaranteed (related to the Riesz Representation Theorem applied to the derivative, which is a linear functional). Our task is to find an explicit expression for this vector.

## 4 The Core Calculation: Finding the Directional Derivative

Let's compute  $D_{\mathbf{h}}f(\mathbf{x})$  for  $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle$ .

1. **Expand  $f(\mathbf{x} + t\mathbf{h})$ :** We substitute  $\mathbf{x} + t\mathbf{h}$  into the definition of  $f$ :

$$\begin{aligned} f(\mathbf{x} + t\mathbf{h}) &= \langle \mathbf{x} + t\mathbf{h}, \mathbf{T}(\mathbf{x} + t\mathbf{h}) \rangle \\ &= \langle \mathbf{x} + t\mathbf{h}, \mathbf{T}\mathbf{x} + t\mathbf{T}\mathbf{h} \rangle && \text{(Linearity of } \mathbf{T} \text{)} \\ &= \langle \mathbf{x}, \mathbf{T}\mathbf{x} + t\mathbf{T}\mathbf{h} \rangle + \langle t\mathbf{h}, \mathbf{T}\mathbf{x} + t\mathbf{T}\mathbf{h} \rangle && \text{(Linearity in 1st arg)} \\ &= \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle + \langle \mathbf{x}, t\mathbf{T}\mathbf{h} \rangle + \langle t\mathbf{h}, \mathbf{T}\mathbf{x} \rangle + \langle t\mathbf{h}, t\mathbf{T}\mathbf{h} \rangle && \text{(Linearity in 2nd arg)} \\ &= \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle + t\langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle + t\langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle + t^2\langle \mathbf{h}, \mathbf{T}\mathbf{h} \rangle && \text{(Linearity of inner product)} \\ &= f(\mathbf{x}) + t(\langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle + \langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle) + t^2\langle \mathbf{h}, \mathbf{T}\mathbf{h} \rangle \end{aligned}$$

2. **Form the difference quotient and take the limit:**

$$\begin{aligned} D_{\mathbf{h}}f(\mathbf{x}) &= \lim_{t \rightarrow 0} \frac{f(\mathbf{x} + t\mathbf{h}) - f(\mathbf{x})}{t} \\ &= \lim_{t \rightarrow 0} \frac{[f(\mathbf{x}) + t(\langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle + \langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle) + t^2\langle \mathbf{h}, \mathbf{T}\mathbf{h} \rangle] - f(\mathbf{x})}{t} \\ &= \lim_{t \rightarrow 0} \frac{t(\langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle + \langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle) + t^2\langle \mathbf{h}, \mathbf{T}\mathbf{h} \rangle}{t} \\ &= \lim_{t \rightarrow 0} [\langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle + \langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle + t\langle \mathbf{h}, \mathbf{T}\mathbf{h} \rangle] \\ &= \langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle + \langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle \end{aligned}$$

So, we've found the directional derivative:  $D_{\mathbf{h}}f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle + \langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle$ .

## 5 The Adjoint Operator's Crucial Role

Our result for  $D_{\mathbf{h}}f(\mathbf{x})$  isn't quite in the form  $\langle \text{something}, \mathbf{h} \rangle$  yet. We need to manipulate the terms so that  $\mathbf{h}$  consistently appears as the second argument in the inner product. This is where the adjoint operator comes in.

**Definition 5.1** (Adjoint Operator). For a linear operator  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{V}$  on an inner product space  $\mathcal{V}$ , its adjoint  $\mathbf{T}^* : \mathcal{V} \rightarrow \mathcal{V}$  is the unique linear operator satisfying:

$$\langle \mathbf{u}, \mathbf{T}\mathbf{v} \rangle = \langle \mathbf{T}^*\mathbf{u}, \mathbf{v} \rangle \quad \text{for all } \mathbf{u}, \mathbf{v} \in \mathcal{V}$$

*Remark 5.1.* Think of the adjoint as the operator that allows  $\mathbf{T}$  to "switch sides" within the inner product. In  $\mathbb{R}^n$  with the standard dot product, if  $\mathbf{T}$  is represented by matrix  $A$ , then  $\mathbf{T}^*$  is represented by the transpose  $A^T$ .

Let's apply this to our directional derivative terms:

- The first term is  $\langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle$ . Using the definition of the adjoint (with  $\mathbf{u} = \mathbf{x}, \mathbf{v} = \mathbf{h}$ ), we get:

$$\langle \mathbf{x}, \mathbf{T}\mathbf{h} \rangle = \langle \mathbf{T}^*\mathbf{x}, \mathbf{h} \rangle$$

This is now in the desired form.

- The second term is  $\langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle$ . We use the symmetry of the real inner product:

$$\langle \mathbf{h}, \mathbf{T}\mathbf{x} \rangle = \langle \mathbf{T}\mathbf{x}, \mathbf{h} \rangle$$

This is also now in the desired form.

## 6 Identifying the Gradient

Substitute these back into the expression for  $D_{\mathbf{h}}f(\mathbf{x})$ :

$$\begin{aligned} D_{\mathbf{h}}f(\mathbf{x}) &= \langle \mathbf{T}^*\mathbf{x}, \mathbf{h} \rangle + \langle \mathbf{T}\mathbf{x}, \mathbf{h} \rangle \\ &= \langle \mathbf{T}^*\mathbf{x} + \mathbf{T}\mathbf{x}, \mathbf{h} \rangle && \text{(Linearity of inner product in 1st arg)} \\ &= \langle (\mathbf{T}^* + \mathbf{T})\mathbf{x}, \mathbf{h} \rangle && \text{(Definition of sum of operators)} \end{aligned}$$

We have successfully expressed the directional derivative in the form  $D_{\mathbf{h}}f(\mathbf{x}) = \langle \text{Vector}, \mathbf{h} \rangle$ . By the definition of the gradient, that "Vector" must be  $\nabla f(\mathbf{x})$ .

**Theorem 6.1** (Gradient of a Quadratic Form). Let  $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle$  be a quadratic form on a real inner product space  $\mathcal{V}$ , where  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{V}$  is a linear operator. The gradient of  $f$  is given by:

$$\nabla f(\mathbf{x}) = (\mathbf{T}^* + \mathbf{T})\mathbf{x}$$

## 7 The Special Case: Self-Adjoint Operators

A very common and important situation is when the operator  $\mathbf{T}$  is self-adjoint.

**Definition 7.1** (Self-Adjoint Operator). An operator  $\mathbf{T}$  is self-adjoint if it equals its adjoint:  $\mathbf{T}^* = \mathbf{T}$ .

*Remark 7.1.* In  $\mathbb{R}^n$  with the dot product, this corresponds to a symmetric matrix ( $A^T = A$ ). Self-adjoint operators have many nice properties, such as real eigenvalues and orthogonal eigenvectors.

If  $\mathbf{T}$  is self-adjoint, our gradient formula simplifies elegantly:

**Corollary 7.1** (Gradient for Self-Adjoint  $\mathbf{T}$ ). If  $\mathbf{T}$  is self-adjoint ( $\mathbf{T}^* = \mathbf{T}$ ), then the gradient of  $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle$  is:

$$\nabla f(\mathbf{x}) = (\mathbf{T} + \mathbf{T})\mathbf{x} = 2\mathbf{T}\mathbf{x}$$

This perfectly matches the familiar result  $\frac{\partial(\mathbf{x}^T A \mathbf{x})}{\partial \mathbf{x}} = 2A\mathbf{x}$  when  $A$  is symmetric.

## 8 Conclusion

By stepping back from coordinates and using the abstract language of inner products and linear operators, we derived the gradient of the quadratic form  $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{T}\mathbf{x} \rangle$  as  $\nabla f(\mathbf{x}) = (\mathbf{T}^* + \mathbf{T})\mathbf{x}$ . This derivation highlights the fundamental role of the inner product structure (defining the gradient and the adjoint) and the linearity of the operator. It also naturally reveals why the transpose (adjoint) appears in the general formula and why it simplifies for symmetric (self-adjoint) operators. This perspective provides a robust foundation for understanding differentiation in more complex vector spaces encountered in advanced mathematics, physics, and engineering.

Keep exploring the connections between abstract concepts and concrete examples – it's where the deepest understanding often lies!