

Boundedness of the Core Inverse Matrix in DeltaNet: Elements Always Lie in $[-1, 1]$

Jianlin Su

2026-01-26

Abstract

From The Brief History of Linear Attention: From Imitation to Innovation and Back we see that the parallel form of DeltaNet involves an inverse matrix of the form $(\mathbf{I} + \mathbf{K}\mathbf{K}^\top \odot \mathbf{M}^-)^{-1}$. Recently, the reader @Arch123 observed experimentally that all elements of this inverse matrix lie in $[-1, 1]$, and asked whether this can be proven mathematically. In this article, we provide two different rigorous proofs for this observation.

1 Problem Description

We begin by restating the problem precisely. Let $\mathbf{K} = [\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_n]^\top \in \mathbb{R}^{n \times d}$, where each $\mathbf{k}_i \in \mathbb{R}^{d \times 1}$ is a column vector with norm not exceeding 1. Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a lower triangular mask matrix defined as

$$M_{i,j} = \begin{cases} 1, & i \geq j \\ 0, & i < j \end{cases}$$

Let $\mathbf{I}$ be the identity matrix, and define $\mathbf{M}^- = \mathbf{M} - \mathbf{I}$. We aim to prove that

$$(\mathbf{I} + \mathbf{K}\mathbf{K}^\top \odot \mathbf{M}^-)^{-1} \in [-1, 1]^{n \times n}.$$

Define $\mathbf{X} = \mathbf{I} + \mathbf{K}\mathbf{K}^\top \odot \mathbf{M}^-$, and let its inverse be $\mathbf{Y} = \mathbf{X}^{-1}$. Writing $\mathbf{X}$ explicitly, we have:

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \mathbf{k}_1^\top \mathbf{k}_2 & 1 & 0 & \cdots & 0 & 0 \\ \mathbf{k}_1^\top \mathbf{k}_3 & \mathbf{k}_2^\top \mathbf{k}_3 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{k}_1^\top \mathbf{k}_{n-1} & \mathbf{k}_2^\top \mathbf{k}_{n-1} & \mathbf{k}_3^\top \mathbf{k}_{n-1} & \cdots & 1 & 0 \\ \mathbf{k}_1^\top \mathbf{k}_n & \mathbf{k}_2^\top \mathbf{k}_n & \mathbf{k}_3^\top \mathbf{k}_n & \cdots & \mathbf{k}_{n-1}^\top \mathbf{k}_n & 1 \end{pmatrix}.$$

Since each $\mathbf{k}_i$ has norm not exceeding 1, all elements of $\mathbf{X}$ lie in $[-1, 1]$. We aim to prove that all elements of $\mathbf{Y} = \mathbf{X}^{-1}$ also lie in $[-1, 1]$.

2 Mathematical Induction

This section presents the first proof using mathematical induction. We note that $\mathbf{X}$ is a lower triangular matrix with ones on the diagonal. This structure allows us to use induction on n .

2.1 Diagonal Properties

We first observe that for lower triangular matrices, the inverse preserves certain diagonal properties. Specifically, if $\mathbf{X}$ is lower triangular with ones on the diagonal, then the submatrices $\mathbf{Y}_{[n-1;n-1]}$ and $\mathbf{Y}_{[1:1]}$ are the inverses of the corresponding submatrices of $\mathbf{X}$. This allows us to assume by induction that all elements of $\mathbf{Y}$ except possibly $Y_{n,1}$ lie in $[-1, 1]$.

2.2 Adjugate Matrix

To prove $Y_{n,1} \in [-1, 1]$, we use the adjugate matrix representation of the inverse. Since $\mathbf{X}$ is lower triangular with ones on the diagonal, its determinant is 1. Thus,

$$Y_{n,1} = (-1)^{n+1} \det \begin{pmatrix} \mathbf{k}_1^\top \mathbf{k}_2 & 1 & 0 & \cdots & 0 \\ \mathbf{k}_1^\top \mathbf{k}_3 & \mathbf{k}_2^\top \mathbf{k}_3 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{k}_1^\top \mathbf{k}_{n-1} & \mathbf{k}_2^\top \mathbf{k}_{n-1} & \mathbf{k}_3^\top \mathbf{k}_{n-1} & \cdots & 1 \\ \mathbf{k}_1^\top \mathbf{k}_n & \mathbf{k}_2^\top \mathbf{k}_n & \mathbf{k}_3^\top \mathbf{k}_n & \cdots & \mathbf{k}_{n-1}^\top \mathbf{k}_n \end{pmatrix}.$$

2.3 Product Form

Computing small cases manually, we observe a pattern:

$$Y_{n,1} = -\mathbf{k}_1^\top (\mathbf{I} - \mathbf{k}_2 \mathbf{k}_2^\top) (\mathbf{I} - \mathbf{k}_3 \mathbf{k}_3^\top) \cdots (\mathbf{I} - \mathbf{k}_{n-1} \mathbf{k}_{n-1}^\top) \mathbf{k}_n.$$

By induction, we prove this formula. Then, using the fact that $\|\mathbf{k}_i\| \leq 1$, we bound:

$$|Y_{n,1}| \leq \|\mathbf{k}_1\| \times \|(\mathbf{I} - \mathbf{k}_2 \mathbf{k}_2^\top) \cdots (\mathbf{I} - \mathbf{k}_{n-1} \mathbf{k}_{n-1}^\top) \mathbf{k}_n\| \leq \|\mathbf{k}_1\| \times \|\mathbf{k}_n\| \leq 1.$$

Thus, $Y_{n,1} \in [-1, 1]$, and by induction, all elements of $\mathbf{Y}$ lie in $[-1, 1]$.

3 Dual Computation

This section presents an alternative proof using the original context of DeltaNet.

3.1 Inverse Form

The recursive form of DeltaNet is:

$$\mathbf{S}_t = \mathbf{S}_{t-1} (\mathbf{I} - \mathbf{k}_t \mathbf{k}_t^\top) + \mathbf{v}_t \mathbf{k}_t^\top = \mathbf{S}_{t-1} + (\mathbf{v}_t - \mathbf{S}_{t-1} \mathbf{k}_t) \mathbf{k}_t^\top.$$

With $\mathbf{S}_0 = \mathbf{0}$, we define $\mathbf{u}_t = \mathbf{v}_t - \mathbf{S}_{t-1} \mathbf{k}_t$, leading to:

$$\mathbf{S}_t = \sum_{i=1}^t \mathbf{u}_i \mathbf{k}_i^\top, \quad \mathbf{u}_t = \mathbf{v}_t - \sum_{i=1}^{t-1} \mathbf{u}_i (\mathbf{k}_i^\top \mathbf{k}_t).$$

Defining $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]^\top$, $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]^\top$, we find:

$$\mathbf{U} = \mathbf{Y} \mathbf{V}, \quad \text{where } \mathbf{Y} = (\mathbf{I} + \mathbf{K} \mathbf{K}^\top \odot \mathbf{M}^-)^{-1}.$$

3.2 Recursive Expansion

Expanding recursively, we derive:

$$\mathbf{S}_t = \sum_{i=1}^t \mathbf{v}_i \mathbf{k}_i^\top \mathbf{H}_{i+1 \rightarrow t},$$

where $\mathbf{H}_{r \rightarrow t} = (\mathbf{I} - \mathbf{k}_r \mathbf{k}_r^\top) \cdots (\mathbf{I} - \mathbf{k}_t \mathbf{k}_t^\top)$, and $\mathbf{H}_{t+1 \rightarrow t} = \mathbf{I}$.

3.3 Comparison of Results

From the two forms, we deduce:

$$Y_{t,i} = \begin{cases} 0, & i > t \\ 1, & i = t \\ -\mathbf{k}_i^\top \mathbf{H}_{i+1 \rightarrow t} \mathbf{k}_t, & i < t \end{cases}$$

To prove $|Y_{t,i}| \leq 1$, we use:

$$|\mathbf{k}_i^\top \mathbf{H}_{i+1 \rightarrow t} \mathbf{k}_t| \leq \|\mathbf{k}_i\| \times \|\mathbf{H}_{i+1 \rightarrow t} \mathbf{k}_t\| \leq \|\mathbf{k}_i\| \times \|\mathbf{k}_t\| \leq 1.$$

4 Summary

We have provided two rigorous proofs that all elements of the inverse matrix $(\mathbf{I} + \mathbf{K} \mathbf{K}^\top \odot \mathbf{M}^-)^{-1}$ in DeltaNet lie in $[-1, 1]$. The first proof uses mathematical induction and determinant properties, while the second leverages the original recursive formulation of DeltaNet and its parallel computation structure.

References

- [1] The Brief History of Linear Attention: From Imitation to Innovation and Back
- [2] Adjugate Matrix (Wikipedia)
- [3] Efficient Inversion of "Diagonal+Low-Rank" Triangular Matrices