Matrix derivatives integral

Derivatives

Vector-by-scalar

Given:

Given vector = and scalar

The derivative of to is a vector where each elem is derivative of elem in to .

=

Scalar-by-Vector

Given:

Given scalar and vector = .

The derivative of to is a vector where each elem is derivative of to corresponding elem .

=

Vector-by-Vector (Jacobian matrix)

Given:

Given vector = and vector =

The derivative of to is a matrix where consist of column vectors with derivative by vector to scalar (see vector-by-scalar subsection). While, the resultant matrix is called

Jacobian matrix.

=

where

=

for all and

Matrix-by-scalar

Given:

Given matrix with size and scalar .

The derivative of to is a matrix where each elem is derivative of corresponding elem in to

=

where

=

Scalar-by-matrix (tangent matrix)

Given:

Given scalar and matrix with size .

The derivative of to is a matrix where each elem is derivative of to each elem in

. The resultant matrix is called tangent matrix.

=

where

=

matrix-by-matrix

Same as below.

vector-by-matrix

Same as below.

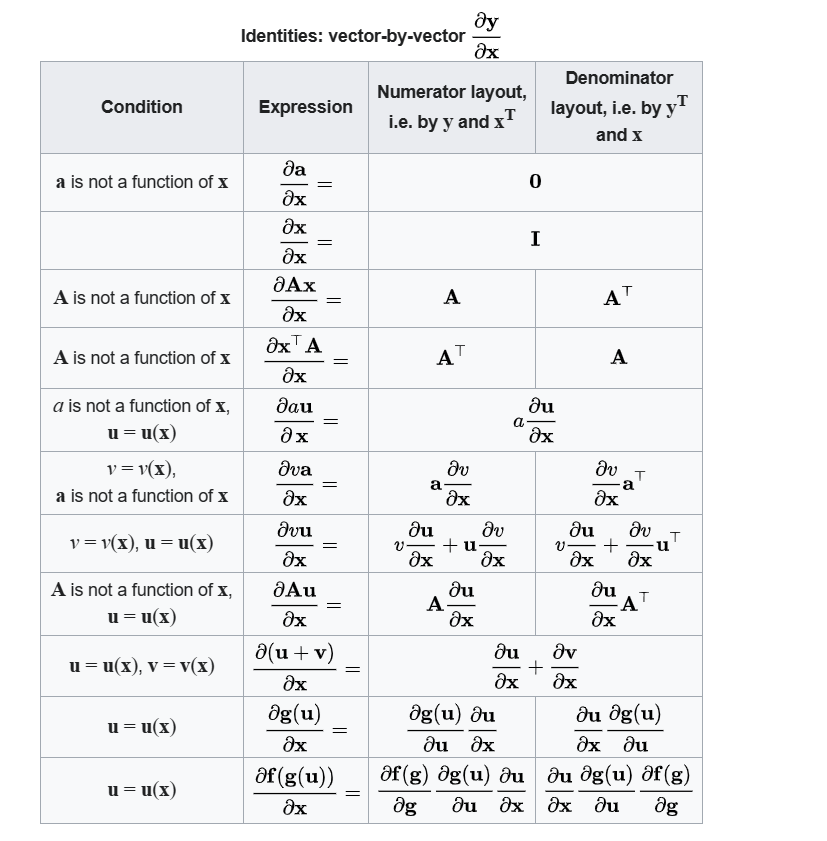
matrix-by-vector

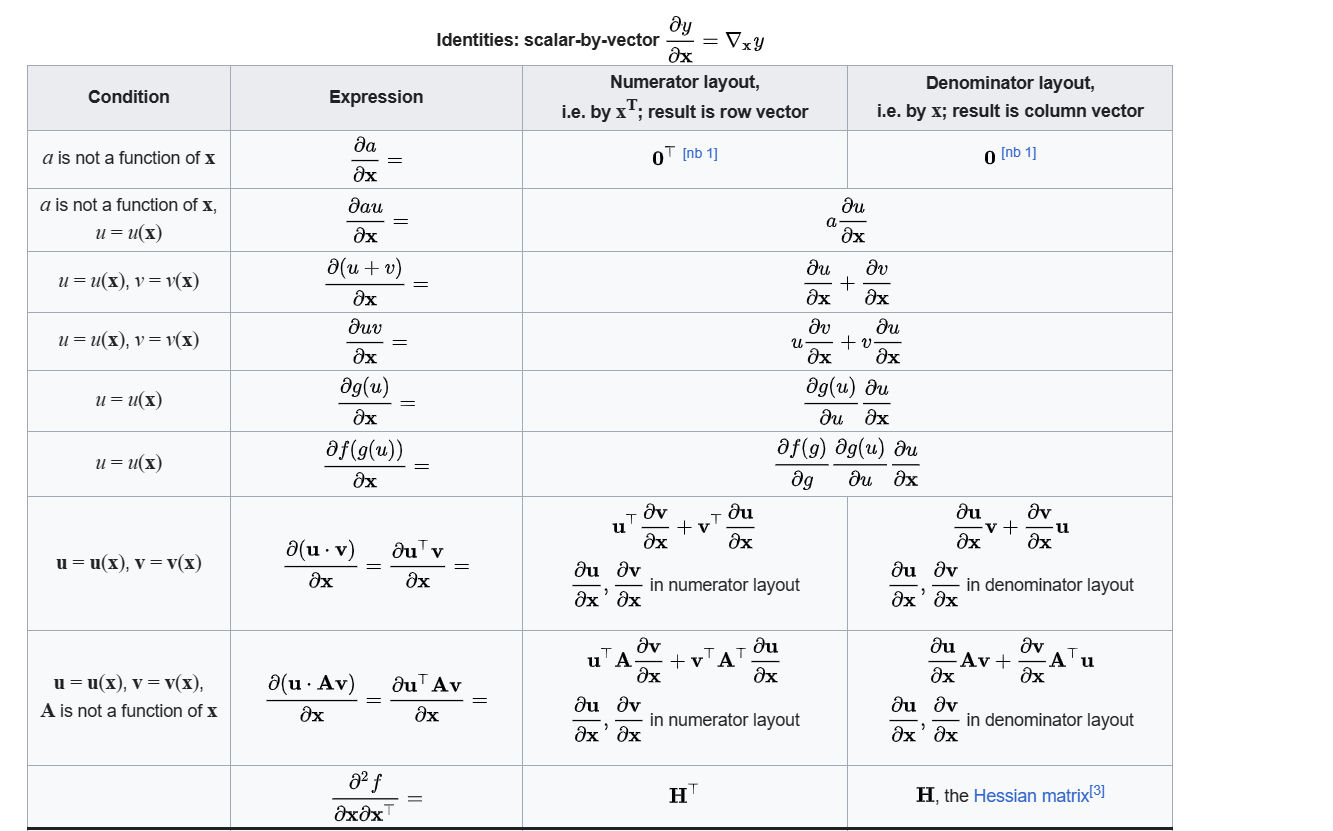
There are many different definition. It has NOT agreed upon yet.

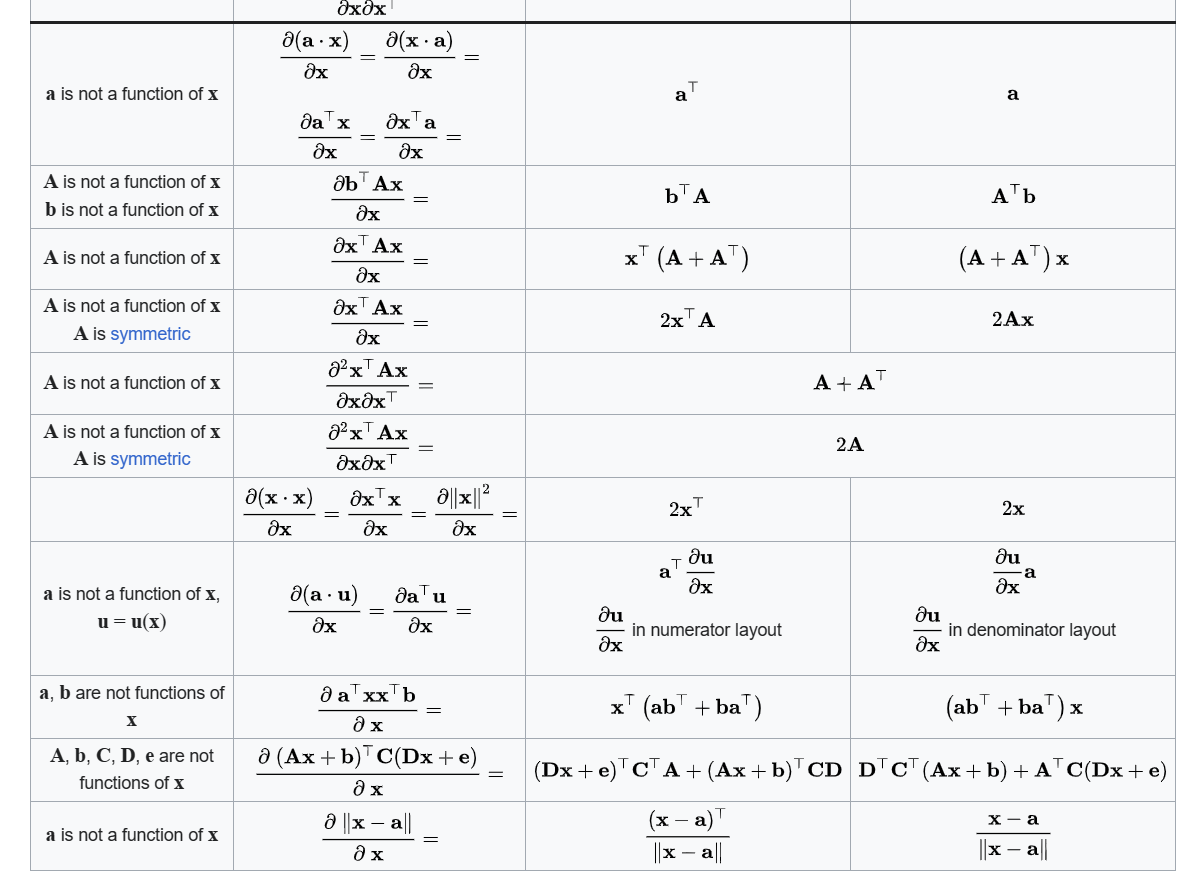
Property

derivative form

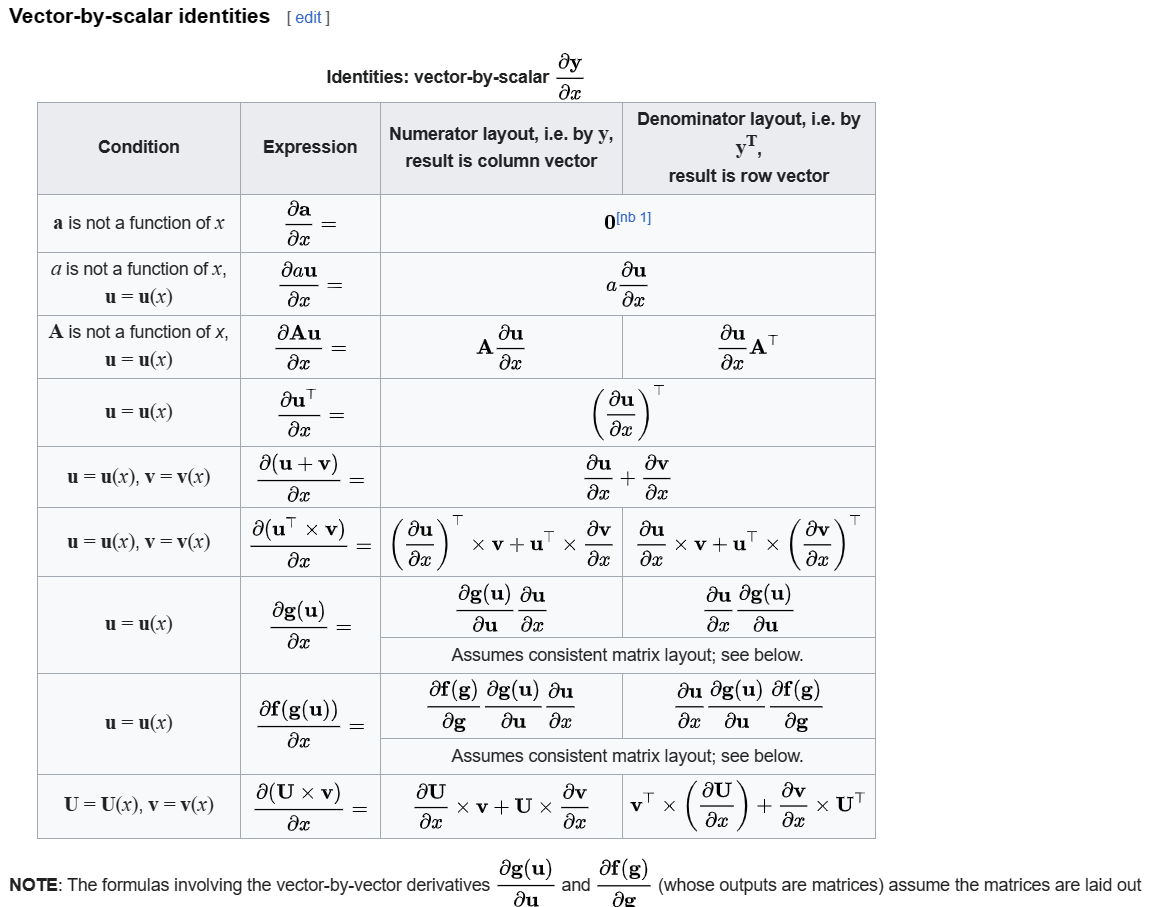
vector to vector



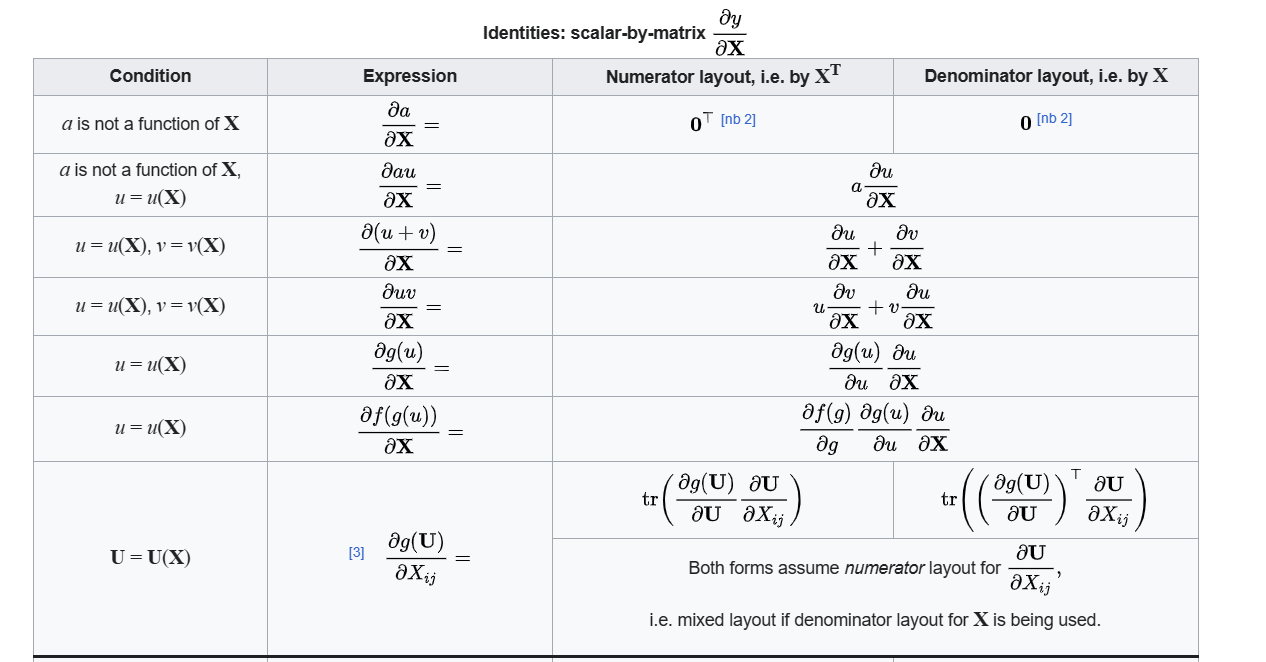




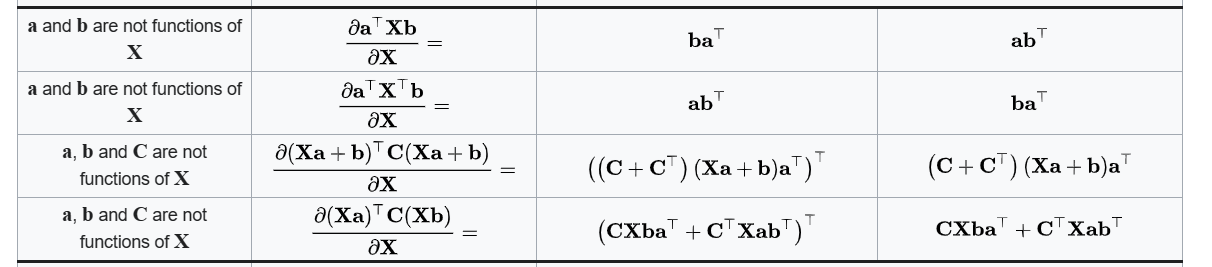
vector to scalar

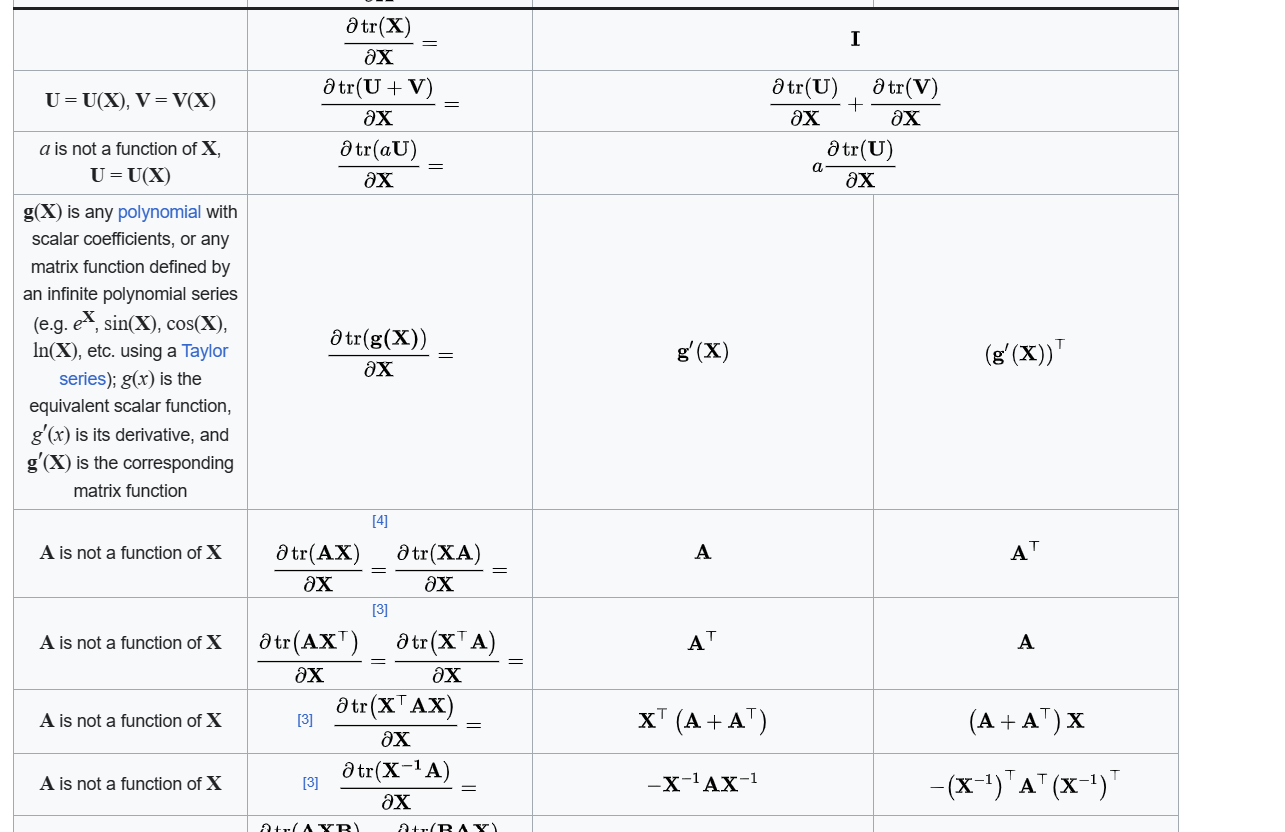


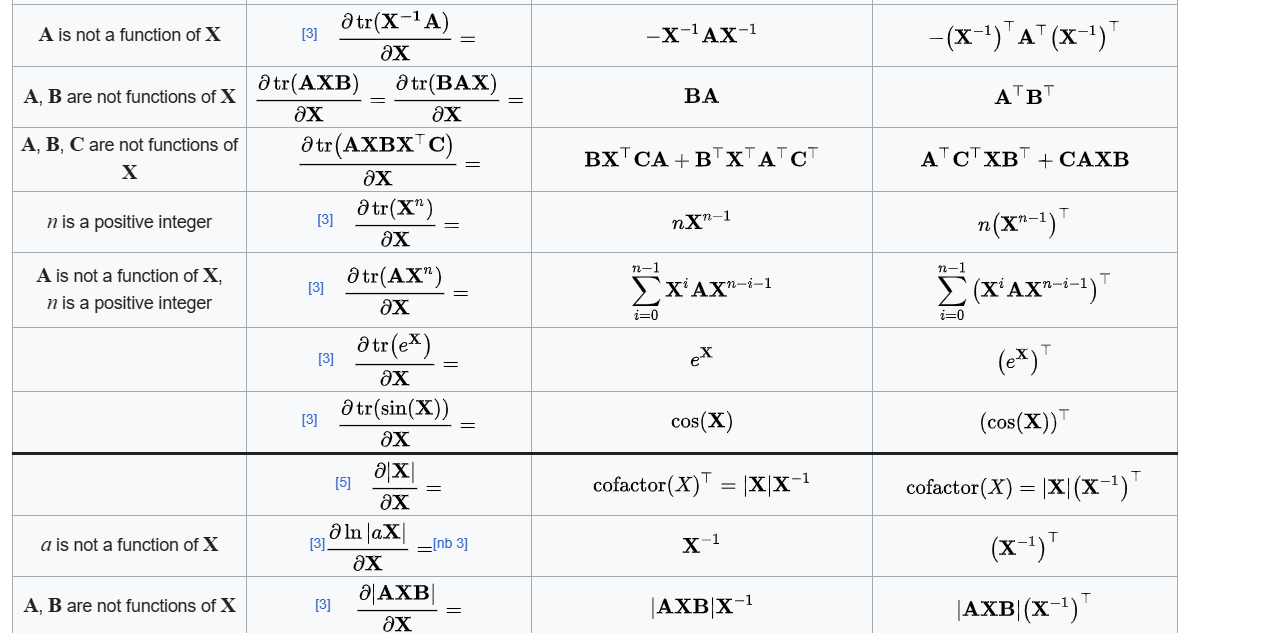
scalar to matrix

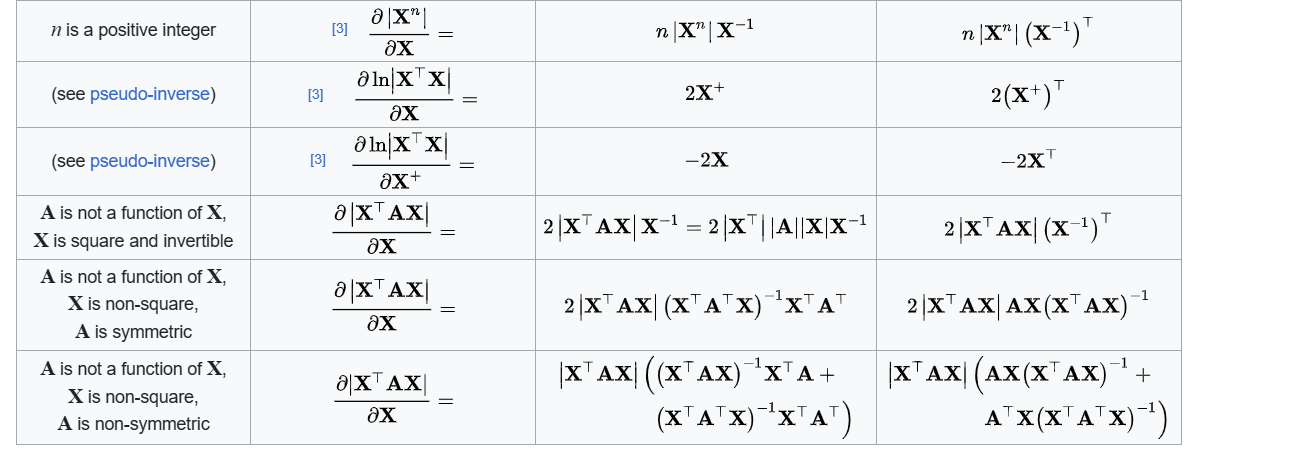


unstructured matrix to matrix

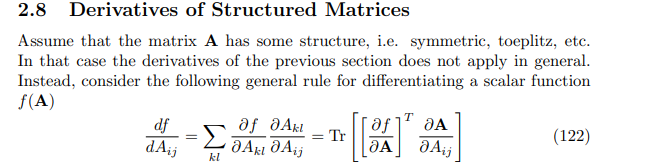




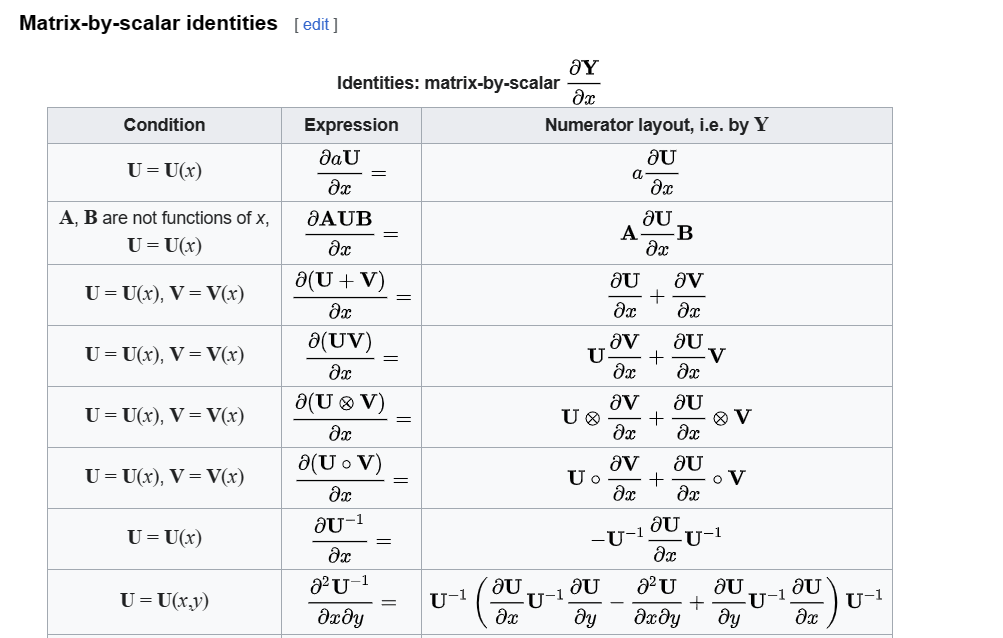


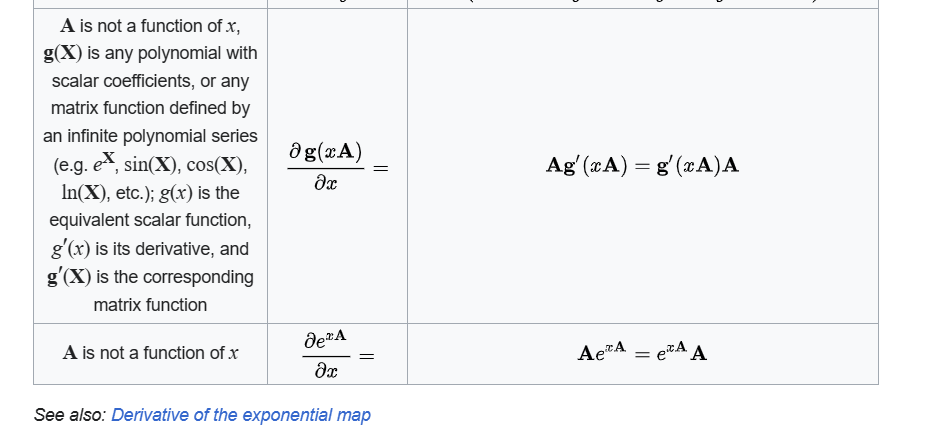


Structured matrix to matrix

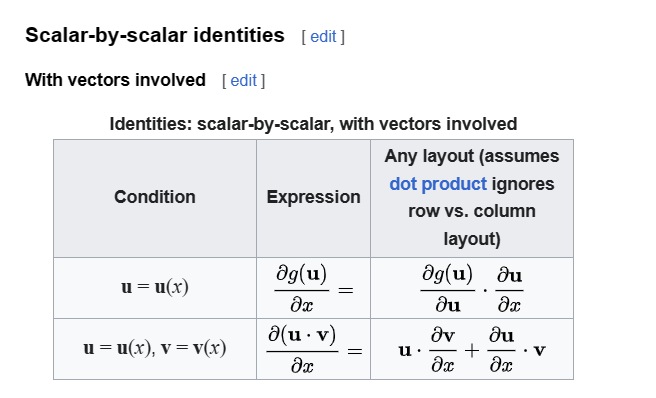


Matrix by scalar

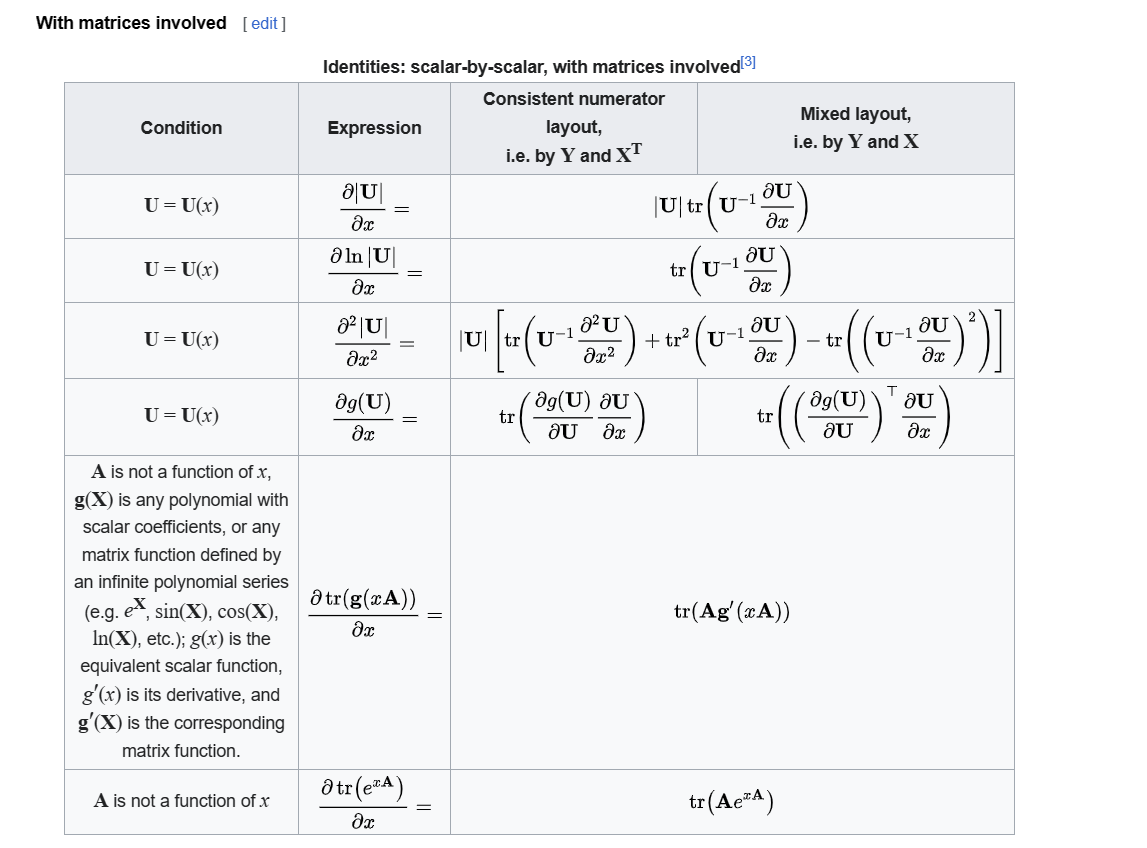




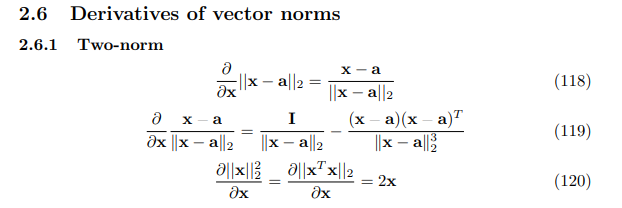
Scalar to scalar (with vector involved)



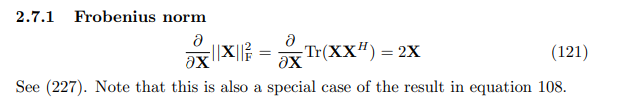
Scalar to scalar (with vector involved)



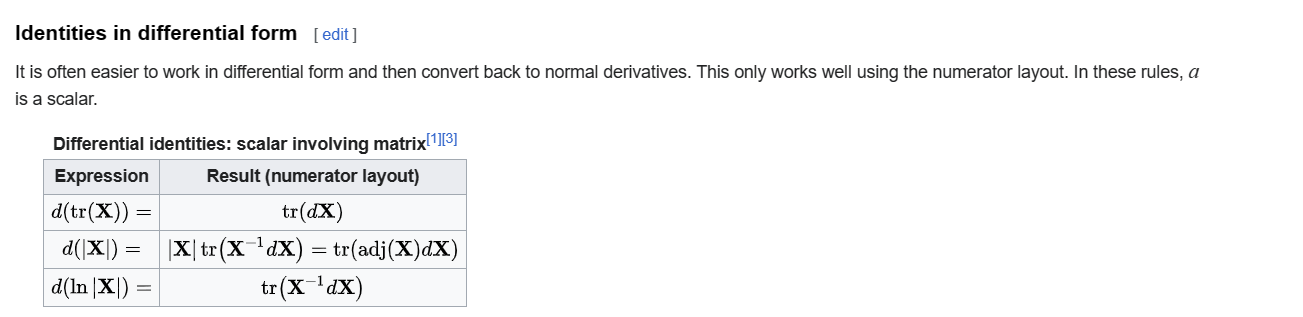
Vector norm to vector

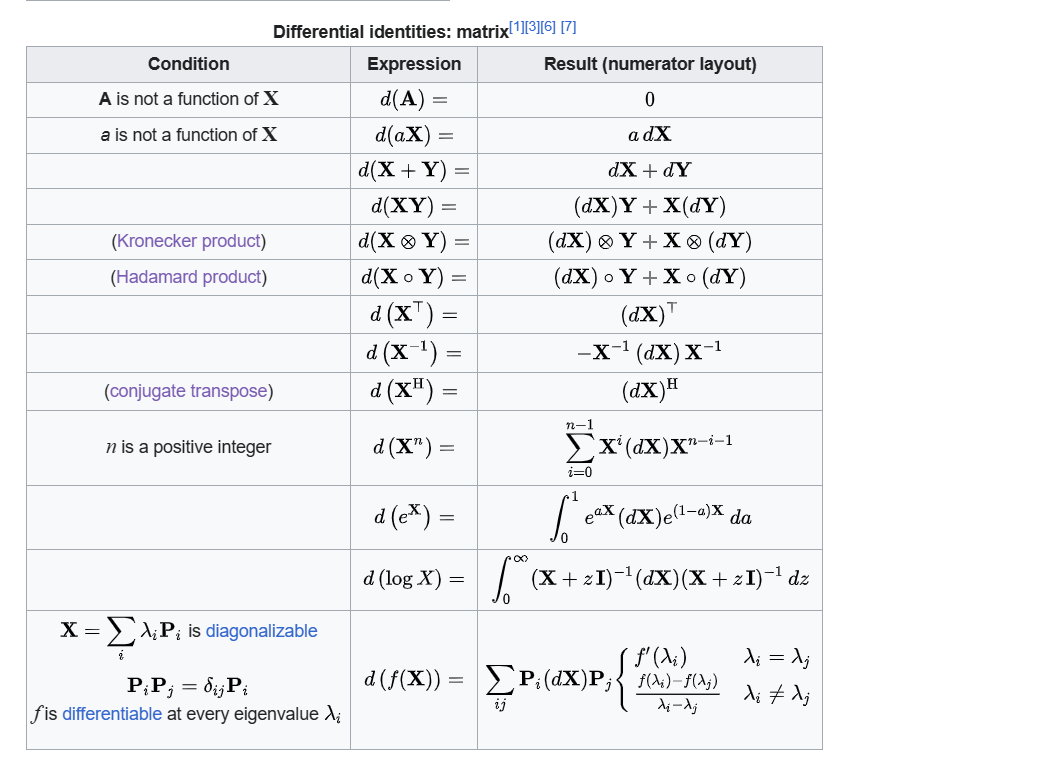


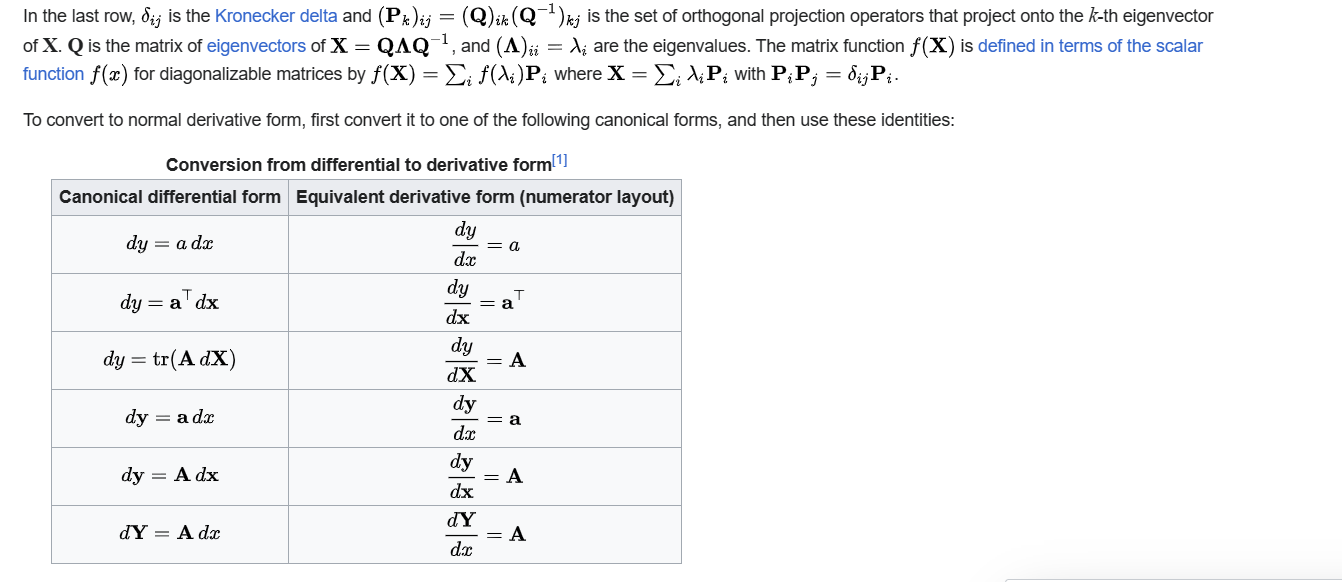
Matrix norm to matrix



differential form



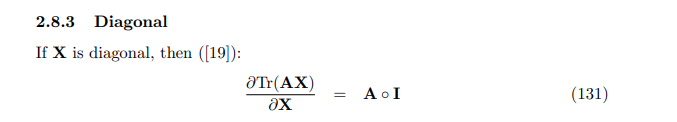




Derivative multiple times

The derivative multiple times refers the derivative it many times using the above approach.

Property about trace



Proof of property

Vector-by-vector





Each elem in be considered as a constant for each elem in . Thus, derivative of it will be zero.





Derivatives of each elem in to itself is 1.





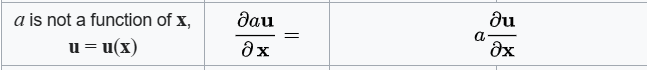
Just expand it then do directly derivatives for .





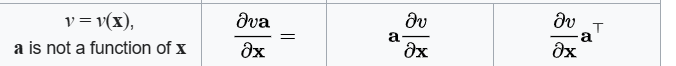
Same as above.





With chain rule.



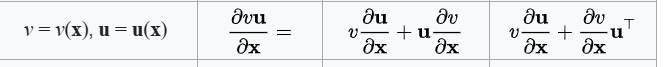


Same as above.



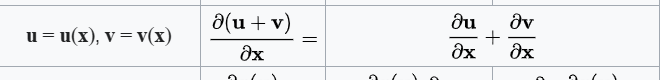






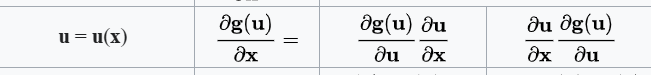
With product rule.





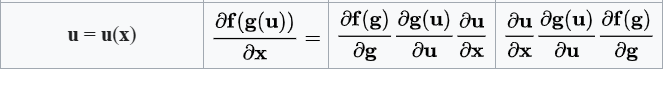
With sum property of derivatives.





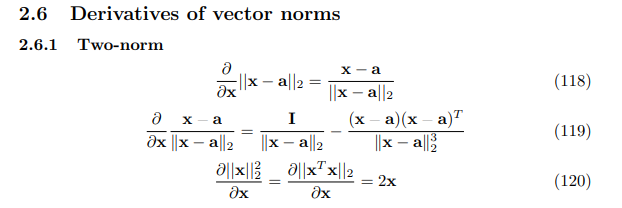
With chain rule.





Same as above.

Vector norm to vector





By p-norm

=

=

=

Evaluate the last term.

=

=

=

Therefore, it will become

=

=



By quotient rule (similar to product rule),

=



=

=

Evaluate the last term.

=

=

=

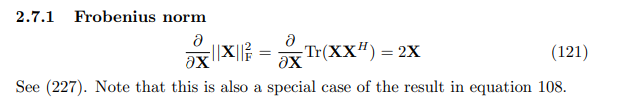
Therefore, it will become

=

=

=

Matrix norm to matrix



=

=

=

=

Scalar-by-vector





Same of proof among these items in Vector-by-vector subsection.





Same of proof among these items in Vector-by-vector subsection.





Same of proof among these items in Vector-by-vector subsection.





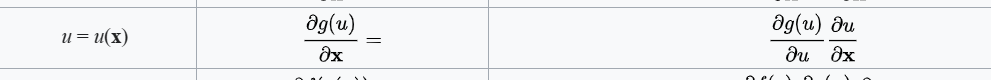
Same of proof among these items in Vector-by-vector subsection.



正在插入影像...

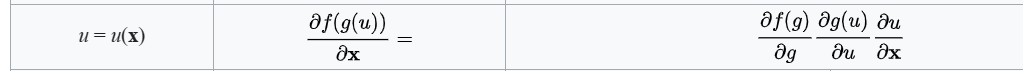
Same of proof among these items in Vector-by-vector subsection.





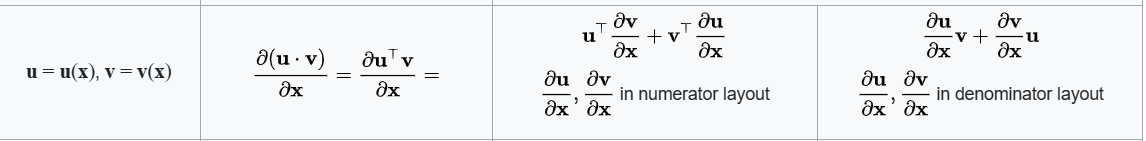
Same of proof among these items in Vector-by-vector subsection.





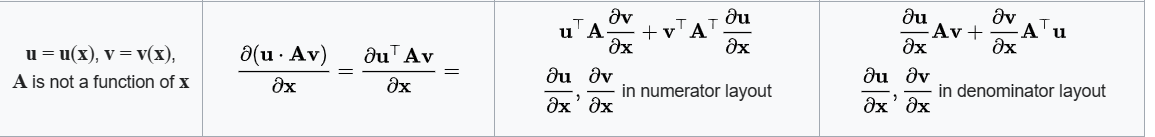
Same of proof among these items in Vector-by-vector subsection.





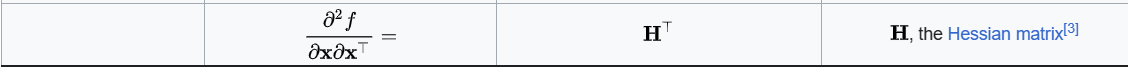
Same of proof among these items in Vector-by-vector subsection.





With chain rule.





By definition of derivatives and Hessian matrix.

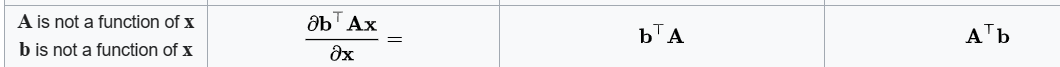
[Hessian matrix - Wikipedia](https://en.wikipedia.org/wiki/Hessian_matrix)





Same of proof among these items in Vector-by-vector subsection.





Same of proof among these items in Vector-by-vector subsection.





=

=

=





By previous property, and definition of symmetric matrix which is =

We have that

=

=

=





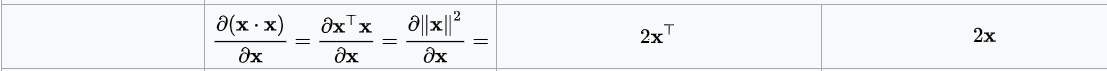
Similar to previous two proof.





With previous proof and definition of symmetric matrix.





By property of inner product,

= where is angle intersection between two vectors

for any vector ,

One has that

= =

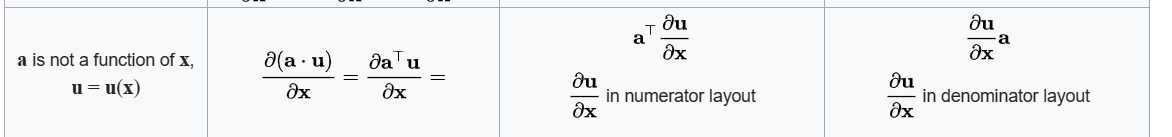
Thus,

=

=

=





With chain rule.

=

=





=

=

=





Let =

=

Then one has that

=

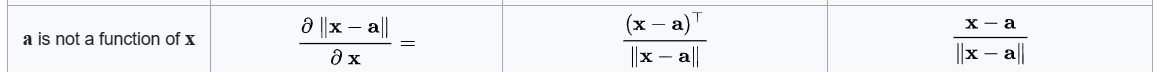
=

=

=

=





Since = ,

one has that

=

And thus, one has that

=

=

=

=

=

Vector-by-scalar





Same of proof among these items in Vector-by-vector subsection.





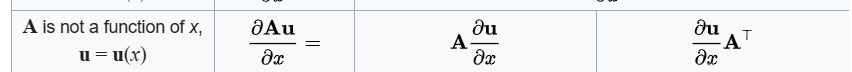
Same of proof among these items in Vector-by-vector subsection.





Same of proof among these items in Vector-by-vector subsection.





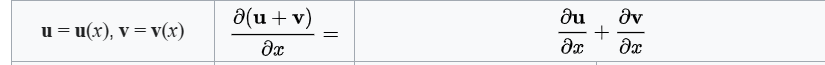
Same of proof among these items in Vector-by-vector subsection.





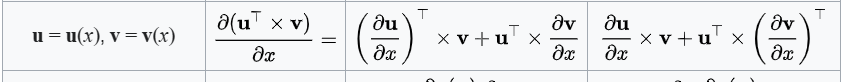
Same of proof among these items in Vector-by-vector subsection.



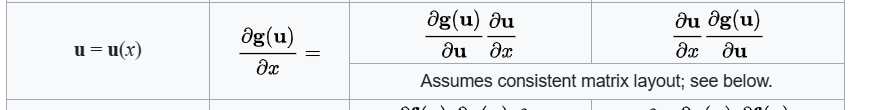


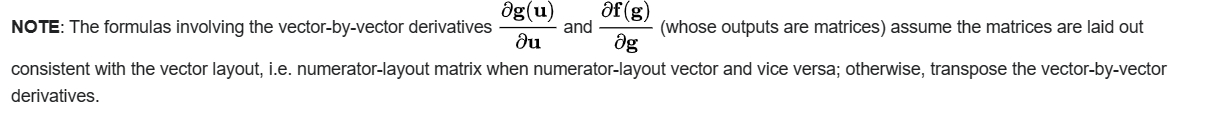
Same of proof among these items in Vector-by-vector subsection.



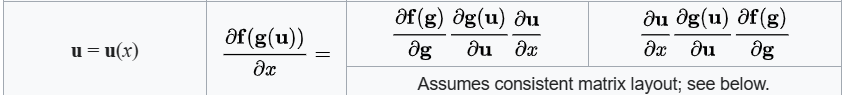


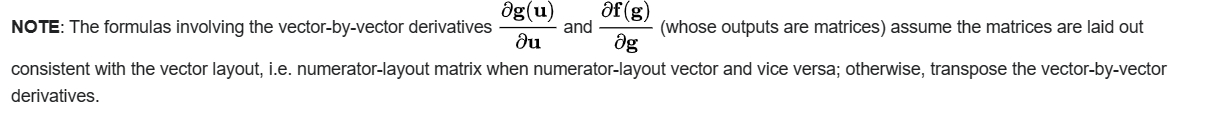






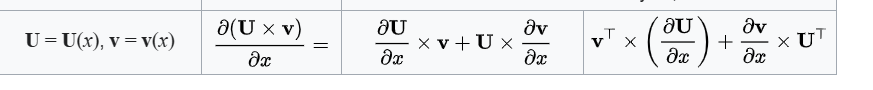






With chain rule.





With product rule.

Scalar-by-matrix





Same of proof among these items in Vector-by-vector subsection.





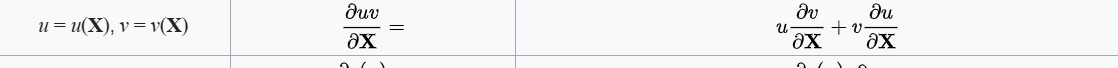
Same of proof among these items in Vector-by-vector subsection.





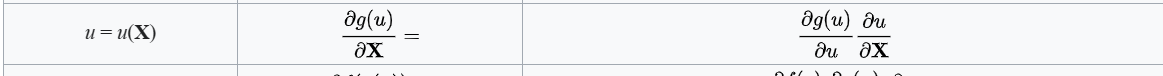
Same of proof among these items in Vector-by-vector subsection.





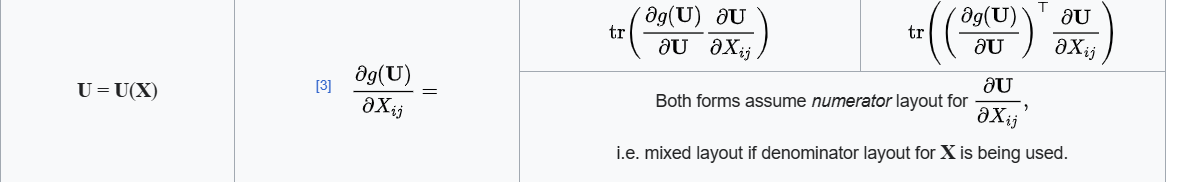
With product rule.





With chain rule.





With chain rule and product rule.



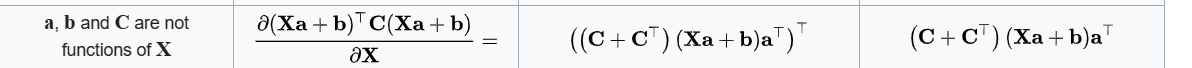






Same of proof among these items in Vector-by-vector subsection.





Similar to proof among these items in Vector-by-vector subsection.

Just replace = . Don’t forget to apply chain rule.

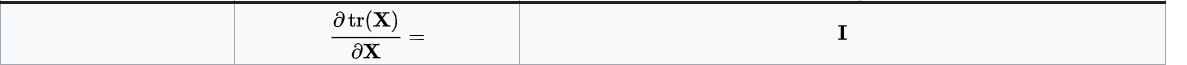




Similar to proof among these items in Vector-by-vector subsection.

Just replace = , = . Don’t forget to apply chain rule.





By property of trace,

= = =

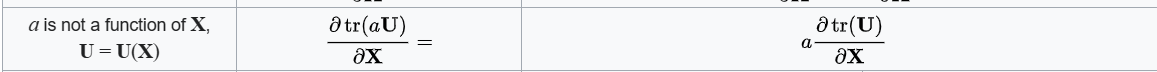




By property about trace

= =

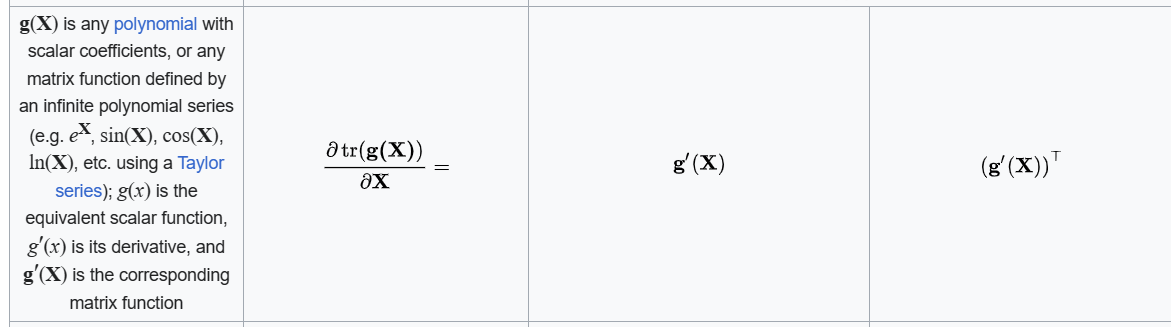




By property of trace,

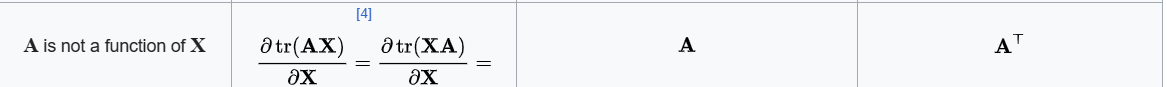
= =





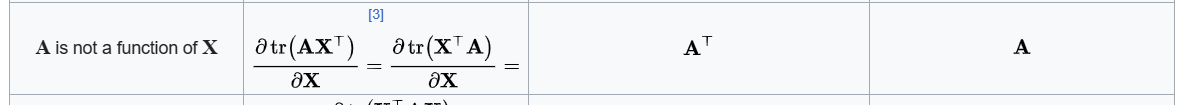
By chain rule.





Since communicativity of multiplication of two matrix in trace.





Similar to above.





Same of proof among these items in Vector-by-vector subsection.



正在插入影像...

First, evaluate

By definition of inverse matrix, one has that

=

Differentiate the equation at both sides. Getting

=

=> =

=> =

=> ==

=> = == =

=

=

=

(by property about trace)

=

(the 1th first by previous claim)

=

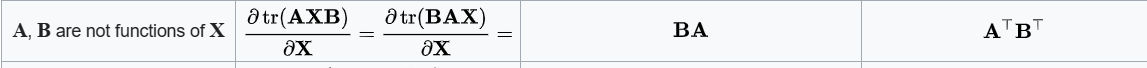
=

=

=

(There are no communicativity for matrix mutliplication in general, however, there are property about cyclic matrix mutliplication in trace function. For more details, see my note trace.docx.)





Again, with property about cyclic matrix mutliplication in trace function.





Again, with property about cyclic matrix mutliplication in trace function.

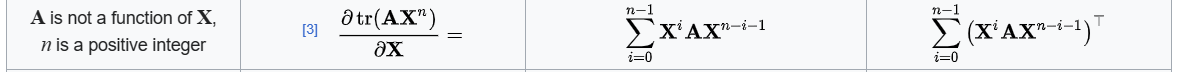


正在插入影像...

=

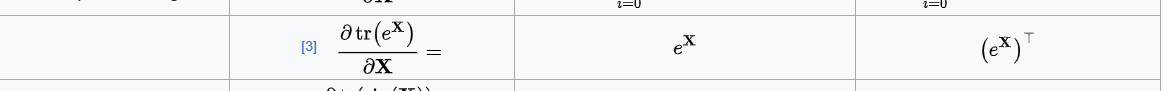
=





Similar as above.





Similar as above.





Similar as above.





By Leinbiz formula of determinant.





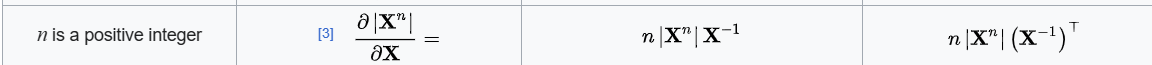
By Leinbiz formula of determinant.





By Leinbiz formula of determinant.





By Leinbiz formula of determinant and chain rule.





With product rule and chain rule.

One has

=

=

=

=

=

(by definition of pseudo-inverse of matrix = )

=

(since ￼ = = (it satisfies all requirements in the prequisite subsection in following section, for more details, see following section) and is NOT equal to which implies that is singular)

=





Prequisites:

Property of

Property of determinant

Relationship between determinant and full rank

Property of derivatives

Property of natural logarithm

Preface:

Since there are NO generalized definition of derivatives of matrix to matrix at present(See the Wiki which is given in Ref section), it is impossible to find (I guess).

While I get stuck in the problem for 2 hours, I found a cookbook given in 3th link in Notes section in Wiki. In section 3.6 of the cookbook, the property about holds. I came up with it, try to proof it. And I found it works for me.

Proof:

Since

1. In this formula, determinant of is used. Thus, has full rank.
2. For any matrix , is a square matrix.

It always holds.

=

which simplifies to

=

=> =

=> =

=> =

=> =

=> =

=> = or

Let =

=

=> =

=>=

Substitute terms with . One has

=

=

=

=

=

=

=

=

=

=

=

=

=

=

=

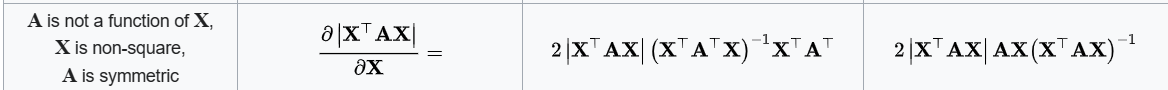
=

=

=

=





With product rule and chain rule.

One has

=

=

=

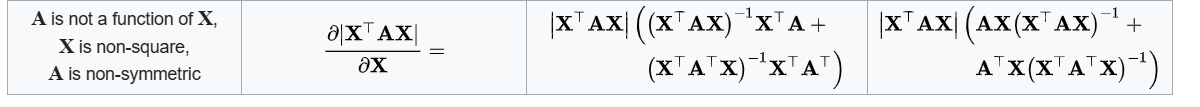
=

=

=

=





Similar to above.

Matrix-by-scalar





Similar to proof among these items in Vector-by-vector subsection.





Similar to proof among these items in Vector-by-vector subsection.





Similar to proof among these items in Vector-by-vector subsection.





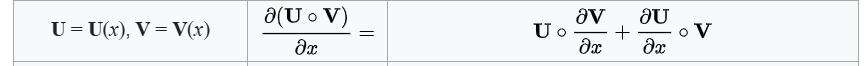
Similar to proof among these items in Vector-by-vector subsection.

正在插入影像...



Similar to proof among these items in Vector-by-vector subsection.

正在插入影像...



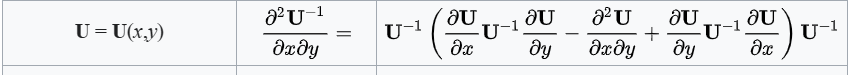
Similar to proof among these items in Vector-by-vector subsection.

正在插入影像...



Similar to proof among these items in Vector-by-vector subsection.

正在插入影像...



=

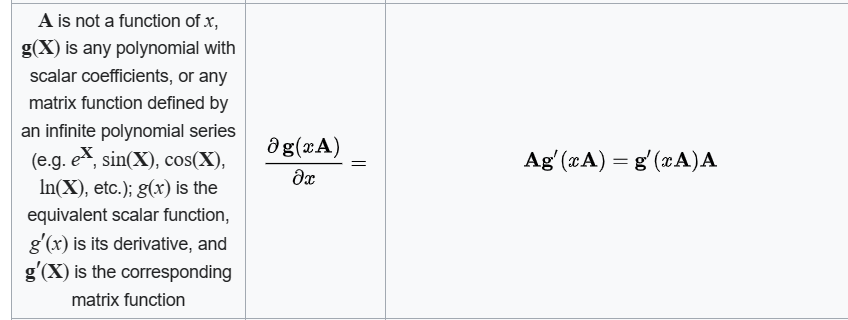
=

= + +

= + +

= +

正在插入影像...



With chain rule.

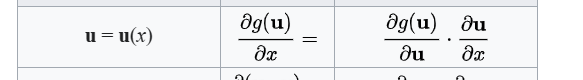
正在插入影像...



With chain rule.

Scalar-by-scalar with vectors involved





Similar to proof among these items in Vector-by-vector subsection.





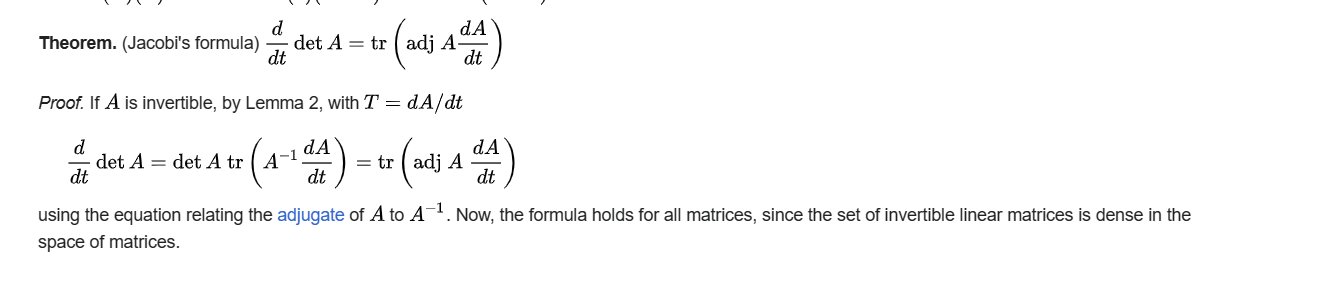
Similar to proof among these items in Vector-by-vector subsection.

Scalar-by-scalar with matrices involved





=



By lemma 2:

From the wiki.

[Jacobi's formula - Wikipedia](https://en.wikipedia.org/wiki/Jacobi%27s_formula)

正在插入影像...



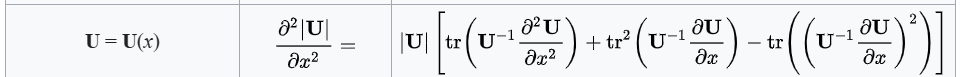
With chain rule.

=

=

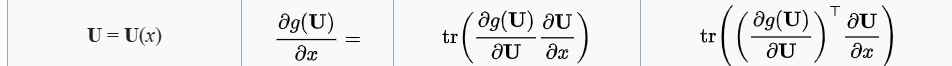
=





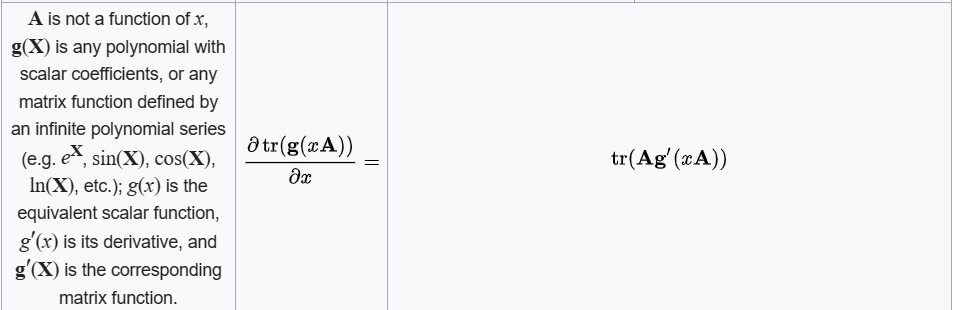
Just differentiate with again.

正在插入影像...



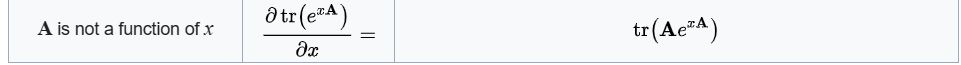
With chain rule and property about trace in matrix.

正在插入影像...



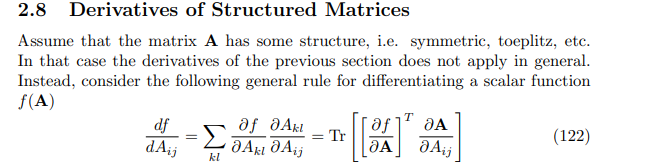
Same as above.

正在插入影像...

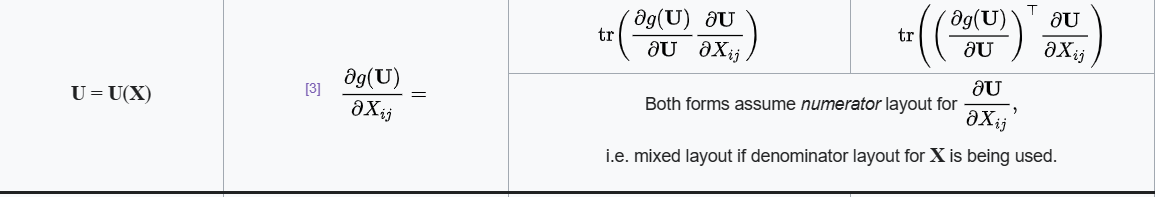


Same above.

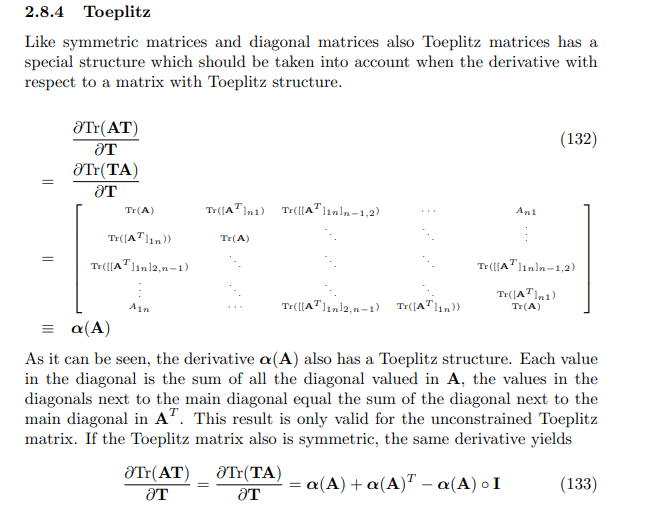
Structured matrix to matrix



Same as the following formula.



Property abot Toeplitz



Differential form

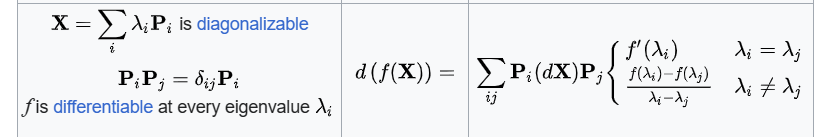
Preface

For differential form, see above proof since it is analytical equivalent but with different representation.

Here, I will only list the items NOT shown above.







Proof of Differential form

(get stuck, not completed yet)



Look at left-hand side and right-hand side of the formula respectively.

On the left-hand side,

=

where

is a shorthand of .

On the right-hand side,

With integration by parts by putting in and in ,

it will be simplified as

=

-

=

-

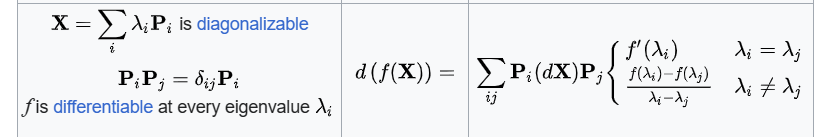
=



On the left-hand side,

=

On the right-hand side,



=

=

=

=

By definition of eigenvalue, one has

Diagonalizable matrix, its eigenvector , and its eigenvalue such that

=

=> =

=> =

=> =

Case 1:

!=

=> =

Suppose there are many eigenvalues represented with =

One has the system

for all and is an integer.

Adding the equations from the system. One has

…

+

=> =

Put the assumption into this equation, it can be written as

=

=> =

=> =

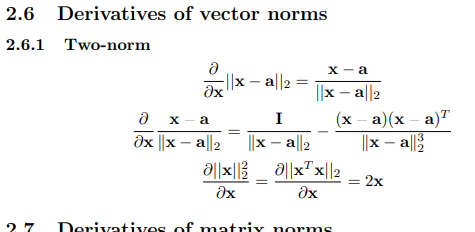
=> =

Case 2:

=

YA!!! can be anything.

Norm in two-dimensional Euclidean space

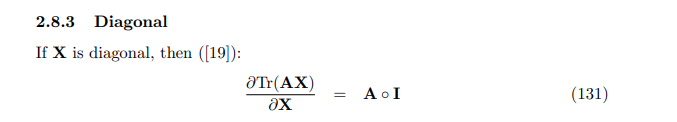


=

=

=

Proof of property about trace



If is diagonal, then

= = , if

= , otherwise

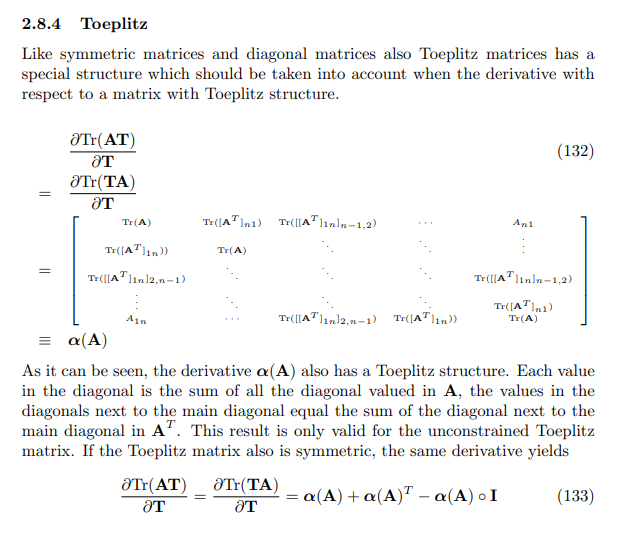
=

=

=

=

Proof of property about Toeplitz



Integral

Rule can be applied to integral.

Performing integral elem-by-elem (each elem in to ,to elem in , each elem in to each elem in )

For more details, see above section.

NOTICE

Notice that in this article, numerator-layout notation is adopted.

Application

Mutlivariate t distribution

Same as below.

Elliptical distribution (special case of multivariate t distribution)

Same as below.

Regression analysis

For more details, see applications section in Wiki.

Ref

[Matrix calculus - Wikipedia](https://en.wikipedia.org/wiki/Matrix_calculus)

[scalar, vector, matrix derivatives](https://en.wikipedia.org/wiki/Diagonalizable_matrix)