# Appendix.01 Linear Algebra

#### Definition.A.1.1 Frobenius norm of the matrix

$$\parallel \mathbf{X} \parallel_F = \sqrt{Tr(X^T X)} = \sqrt{Tr(X X^T)} = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}^2$$

#### Definition.A.1.2 Inner product of vectors

Let vectors $\mathbf{w}$ and $\mathbf{x}$ be $n \times 1$ vector.
$\mathbf{w}^T\mathbf{x}$ is called inner product of vectors.

$$\mathbf{w}^T\mathbf{x} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} w_i x_i$$

#### Definition.A.1.3 Outter product of vectors

Let vectors $\mathbf{w}$ and $\mathbf{x}$ be $n \times 1$ vectors.
$\mathbf{x}\mathbf{w}^T$ is called outter product of vectors.

$$\mathbf{x}\mathbf{w}^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} x_1\mathbf{w}^T \\ x_2\mathbf{w}^T \\ \vdots \\ x_n\mathbf{w}^T \end{bmatrix} = \begin{bmatrix} x_1w_1 & x_1w_2 & \cdots & x_1w_n \\ x_2w_1 & x_2w_2 & \cdots & x_2w_n \\ \vdots & \vdots & \ddots & \vdots \\ x_nw_1 & x_nw_2 & \cdots & x_nw_n \end{bmatrix}$$

**Definition.A.1.4 Matrix-vector multiplication**

Let $W \ : \ m \times n$ and $\mathbf{x} \ : \ n \times 1$

(1)

$$W\mathbf{x} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_n \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} = \sum_{i=1}^n \mathbf{w}_i \mathbf{x}_i$$

(2)

$$W\mathbf{x} = \begin{bmatrix} \bar{\mathbf{w}}_1{}^T \\ \bar{\mathbf{w}}_2{}^T \\ \vdots \\ \bar{\mathbf{w}}_m{}^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} \bar{\mathbf{w}}_1{}^T\mathbf{x} \\ \bar{\mathbf{w}}_2{}^T\mathbf{x} \\ \vdots \\ \bar{\mathbf{w}}_m{}^T\mathbf{x} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n w_{1i}x_i \\ \sum_{i=1}^n w_{2i}x_i \\ \vdots \\ \sum_{i=1}^n w_{mi}x_i \end{bmatrix} \text{where } \bar{\mathbf{w}}_i = \text{(transpose of i-th row vector of } W)$$

(3)

$$W\mathbf{x} = \begin{bmatrix} W_1 & W_2 & \cdots & W_N \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \cdots \\ \mathbf{x}_N \end{bmatrix} = \sum_{i=1}^N W_i \mathbf{x}_i \text{ where } \sum_{i=1}^N n_i = n$$

(4)

$$W\mathbf{x} = \begin{bmatrix} \bar{W_1}^T \\ \bar{W_2}^T \\ \vdots \\ \bar{W_M}^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} \bar{W_1}^T\mathbf{x} \\ \bar{W_2}^T\mathbf{x} \\ \vdots \\ \bar{W_M}^T\mathbf{x} \end{bmatrix} \text{where } \sum_{i=1}^M m_i = m$$

**Definition.A.1.5 Matrix-matrix multiplication**

Let $W\ :\ m\times n$ and $X\ :\ n\times l$.

(1)

$$WX = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_n \end{bmatrix} \begin{bmatrix} \bar{\mathbf{x}}_1{}^T \\ \bar{\mathbf{x}}_2{}^T \\ \cdots \\ \bar{\mathbf{x}}_n{}^T \end{bmatrix} = \sum_{i=1}^{n} \mathbf{w}_i \bar{\mathbf{x}}_i{}^T$$

(2)

$$WX = \begin{bmatrix} \bar{\mathbf{w}}_1{}^T \\ \bar{\mathbf{w}}_2{}^T \\ \vdots \\ \bar{\mathbf{w}}_m{}^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_l \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{w}}_1{}^T\mathbf{x}_1 & \bar{\mathbf{w}}_1{}^T\mathbf{x}_2 & \cdots & \bar{\mathbf{w}}_1{}^T\mathbf{x}_l \\ \bar{\mathbf{w}}_2{}^T\mathbf{x}_1 & \bar{\mathbf{w}}_2{}^T\mathbf{x}_2 & \cdots & \bar{\mathbf{w}}_2{}^T\mathbf{x}_l \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathbf{w}}_m{}^T\mathbf{x}_1 & \bar{\mathbf{w}}_m{}^T\mathbf{x}_2 & \cdots & \bar{\mathbf{w}}_m{}^T\mathbf{x}_l \end{bmatrix}$$

(3)

$$WX = \begin{bmatrix} W_1 & W_2 & \cdots & W_N \end{bmatrix} \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_N \end{bmatrix} = \sum_{i=1}^{N} W_i \bar{X}_i \text{ where } W_i\ :\ m\times n_i\ ,\ \ \bar{X}_i\ :\ n_i\times l\ ,\ \ n=\sum_{i=1}^{N} n_i$$

(4)

$$WX = \begin{bmatrix} \bar{W}_1^T \\ \bar{W}_2^T \\ \vdots \\ \bar{W}_M^T \end{bmatrix} \begin{bmatrix} X_1 & X_2 & \cdots & X_L \end{bmatrix} = \begin{bmatrix} \bar{W}_1^T X_1 & \bar{W}_1^T X_2 & \cdots & \bar{W}_1^T X_L \\ \bar{W}_2^T X_1 & \bar{W}_2^T X_2 & \cdots & \bar{W}_2^T X_L \\ \vdots & \vdots & \ddots & \vdots \\ \bar{W}_M^T X_1 & \bar{W}_M^T X_2 & \cdots & \bar{W}_M^T X_L \end{bmatrix} \quad \text{where}$$

$\bar{W}_i\ :\ m_i\times n\ ,\ \ m=\sum_{i=1}^{M} m_i\ ,\ \ X_i\ :\ n\times l_i\ ,\ \ l=\sum_{i=1}^{l} l_i$

**Definition.A.1.6 Quadratic Forms**

In vectors-matrix, let $w \ : \ n \times 1$ , $R \ : \ n \times n$

$$\begin{aligned}
\mathbf{w}^T R \mathbf{w} &= \mathbf{w}^T \sum_{j=1}^{n} \mathbf{r}_j \mathbf{w}_j \\
&= \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \begin{bmatrix} \sum_{j=1}^{n} r_{1j} w_j \\ \sum_{j=1}^{n} r_{2j} w_j \\ \vdots \\ \sum_{j=1}^{n} r_{nj} w_j \end{bmatrix} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} w_i r_{ij} w_j
\end{aligned}$$

In matrix-matrix,

$$W^T R W = \sum_{i=1}^{n} \sum_{j=1}^{n} W_i R_{ij} W_j \text{ where } W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix}, R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{nn} \end{bmatrix}$$

**Definition.A.1.7 Unitary(Orthogonal) matrix**

A $n \times n$ matrix $Q$ is called unitary(orthogonal) matrix if

$$Q^T Q = Q Q^T = I \ where \ Q \ : \ n \times n$$

$(i)$
Let $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$

$$\begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T \mathbf{q}_1 & \mathbf{q}_1^T \mathbf{q}_2 & \cdots & \mathbf{q}_1^T \mathbf{q}_n \\ \mathbf{q}_2^T \mathbf{q}_1 & \mathbf{q}_2^T \mathbf{q}_2 & \cdots & \mathbf{q}_2^T \mathbf{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n^T \mathbf{q}_1 & \mathbf{q}_n^T \mathbf{q}_2 & \cdots & \mathbf{q}_n^T \mathbf{q}_n \end{bmatrix} \text{,which is } \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}$$

**Definition.A.1.8 Eigenvalues and Eigenvectors**

$$\begin{aligned}
\text{If } \exists \ \lambda_i \in \mathbb{R} \ s.t. \ R\mathbf{q} = \lambda \mathbf{q} \quad &\Leftrightarrow \quad (R - \lambda I)\mathbf{q} = \mathbf{0} \\
&\Leftrightarrow \quad P(\lambda) = det(R - \lambda I) = 0
\end{aligned}$$

$where \ R \ : \ n \times n$

$\lambda$ is called eigenvalue and $\mathbf{q}$ is called eigenvector.
Eigenvectors are often normalized such that

$$\| \mathbf{q}_i \| = 1, \ i = 1, 2, \cdots, n$$

**Definition.A.1.9 Symmetric Matrix**

$$R^T = R \ where \ R \ : \ n \times n$$

,which is called symmetric matrix.

$(i)$ The eigenvalues of a symmetric matrix are real-valued, not complex-valued.

$$\lambda_i \in \mathbb{R} \ for \ i = 1, 2, \cdots, n$$

$(ii)$ The eigenvectors of a symmetric matrix are orthonormal.

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

**Definition.A.1.10 Positive definite matrix and positive semi-definite matrix**

A symmetric $n \times n$ real matrix $R$ is said to be positive definite if the scalar $\mathbf{z}^T R \mathbf{z}$ is positive for every non-zero column vector $\mathbf{z}$ of n real numbers. It can be written

$$R \succ \mathbf{0}$$

A symmetric $n \times n$ real matrix $R$ is said to be semi-positive definite if the scalar $\mathbf{z}^T R \mathbf{z}$ isn't negative for every non-zero column vector $\mathbf{z}$ of n real numbers. It can be written

$$R \succeq \mathbf{0}$$

**Theorem.A.1.1 Eigenvalue Decomposition(EVD) and singular value decomposition(SVD)**

Any symmetric matrix $R$ can be decomposed as

$$R = Q\Lambda Q^T = \sum_{i=1}^{n} \lambda_i \mathbf{q}_i \mathbf{q}_i^T$$

$$where\ Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

**Proof.**

$$R\mathbf{q}_i = \lambda_i \mathbf{q}_i,\ \ i = 1, 2, \cdots, n$$

$$Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$$

$$\begin{aligned} RQ &= R\begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{q}_1 & \lambda_2\mathbf{q}_2 & \cdots \lambda_n\mathbf{q}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \\ &= Q\Lambda \quad \Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \end{aligned}$$

$$R = Q\Lambda Q^{-1} = Q\Lambda Q^T \ \ (\because\ Q^{-1} = Q^T) \quad \blacksquare$$

Singular value decomposition(SVD) Any square matrix $A$ can be decomposed as

$$A = V\Sigma U^T$$

$$\text{where } V \text{ is an unitary matrix of eigenvectors of } A^T A \\ , U^T \text{ is an unitary matrix of eigenvectors of } AA^T \ \ .$$

**Theorem.A.1.2 Interpretation of EVD**

Let $\mathbf{x}\ :\ n \times 1$ and $R\ :\ n \times n$ .
$\mathbf{x}^T R\mathbf{x} = 1\ for\ A \succ 0$
that is an ellipse in n-dimensional space.

Axes : eigenvectors $\{\mathbf{q}_i\}_{i=1}^{n}$
Half-Length of each axis : $\frac{1}{\sqrt{\lambda_i}}$

**Proof.**
$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T V\Lambda V^T\mathbf{x} = 1$
$\mathbf{y}^T \Lambda\mathbf{y} = \sum_{j=1}^{n} \lambda_j y_j^2 = 1\ where\ \mathbf{y} = V^T\mathbf{x}$
$\sum_{j=1}^{n} \lambda_j y_j^2$ is a equation for the ellipse in n-dimension. $\blacksquare$

**Definition.A.1.11 Trace**

Let $A$ be a matrix.

$$Tr(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

$(i)\ Tr(c) = c\ for\ some\ scalar\ c$
$(ii)\ Tr(AB) = Tr(BA)$
$(iii)\ Tr(\mathbf{x}\mathbf{x}^T) = Tr(\mathbf{x}^\mathbf{T}\mathbf{x}) = \mathbf{x}^\mathbf{T}\mathbf{x} = \ \| x \|^2\ for\ some\ vector\ \mathbf{x}$
$(vi)\ Tr(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i\ where\ A\ :\ n \times n$

**Proof.**
$(i)\ Definition$
$(ii)\ Trivial \quad \blacksquare$
$(iii)\ Trivial \quad \blacksquare$

$$\begin{aligned}(vi)\ Tr(A) &= Tr(Q\Lambda Q^T) \\ &= Tr(Q^T Q\Lambda)\ (\ \because\ (ii)\ ) \\ &= Tr(\Lambda)\ (\ \because\ Q\ is\ unitary\ ) \\ &= \sum_{i=1}^{n} \lambda_i \quad \blacksquare\end{aligned}$$

**Definition.A.1.12 Gradient of a scalar function with respect to a vector**

Let the function $f\ :\ \mathbb{R}^m \to \mathbb{R}$ be.

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}} = \ \bigtriangledown_\mathbf{w}\ f(\mathbf{w}) = \begin{bmatrix} \frac{\partial f(\mathbf{w})}{\partial w_1} \\ \frac{\partial f(\mathbf{w})}{\partial w_2} \\ \cdots \\ \frac{\partial f(\mathbf{w})}{\partial w_m} \end{bmatrix}$$

**Definition.A.1.13 Gradient of a vector function with respect to a vector**

Let the function $\mathbf{g}\ :\ \mathbb{R}^m \to \mathbb{R}^n$ be.

If $\mathbf{g}(\mathbf{w}) = \begin{bmatrix} g_1(\mathbf{w}) \\ g_2(\mathbf{w}) \\ \vdots \\ g_n(\mathbf{w}) \end{bmatrix}$. and $w\ :\ m \times 1$,

$$\frac{\partial \mathbf{g}(\mathbf{w})}{\partial \mathbf{w}} = \ \bigtriangledown_\mathbf{w}\ \mathbf{g}(\mathbf{w}) = \begin{bmatrix} \frac{\partial g_1(\mathbf{w})}{\partial w_1} & \frac{\partial g_1(\mathbf{w})}{\partial w_2} & \cdots & \frac{\partial g_1(\mathbf{w})}{\partial w_m} \\ \frac{\partial g_2(\mathbf{w})}{\partial w_1} & \frac{\partial g_2(\mathbf{w})}{\partial w_2} & \cdots & \frac{\partial g_2(\mathbf{w})}{\partial w_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m(\mathbf{w})}{\partial w_1} & \frac{\partial g_m(\mathbf{w})}{\partial w_2} & \cdots & \frac{\partial g_m(\mathbf{w})}{\partial w_m} \end{bmatrix}$$

#### Definition.A.1.14 Hessian matrix of a scalar function with respect to a vector

Let the function $f : \mathbb{R}^m \rightarrow \mathbb{R}$ and $\mathbf{w} : m \times 1$ be.

$$\mathbf{H} = \frac{\partial}{\partial \mathbf{w}} \nabla_{\mathbf{w}}^2 f(\mathbf{w}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{w})}{\partial w_1^2} & \frac{\partial^2 f(\mathbf{w})}{\partial w_1 \partial w_2} & \cdots & \frac{\partial^2 f(\mathbf{w})}{\partial w_1 \partial w_m} \\ \frac{\partial^2 f(\mathbf{w})}{\partial w_2 \partial w_1} & \frac{\partial^2 f(\mathbf{w})}{\partial w_2^2} & \cdots & \frac{\partial^2 f(\mathbf{w})}{\partial w_2 \partial w_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{w})}{\partial w_m \partial w_1} & \frac{\partial^2 f(\mathbf{w})}{\partial w_m \partial w_2} & \cdots & \frac{\partial^2 f(\mathbf{w})}{\partial w_m^2} \end{bmatrix}$$

#### Definition.A.1.15 Gradient of a scalar function with respect to a matrix

Let the function $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ be a function.

$$\frac{\partial f(W)}{\partial W} = \nabla_W f(W) = \begin{bmatrix} \frac{\partial f(W)}{\partial w_{11}} & \frac{\partial f(W)}{\partial w_{12}} & \cdots & \frac{\partial f(W)}{\partial w_{1n}} \\ \frac{\partial f(W)}{\partial w_{21}} & \frac{\partial f(W)}{\partial w_{22}} & \cdots & \frac{\partial f(W)}{\partial w_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(W)}{\partial w_{m1}} & \frac{\partial f(W)}{\partial w_{m2}} & \cdots & \frac{\partial f(W)}{\partial w_{mn}} \end{bmatrix}$$

#### Theorem.A.1.3 Geometric solution of distance between vector and hyperplane

Let a vector $\mathbf{x}_0$ and a hyperplane $\mathbf{w}^T\mathbf{x} + b = 0$ be. The distance between a vector $\mathbf{x}_0$ and a hyperplane $\mathbf{w}^T\mathbf{x} + b = 0$ $r$ is

$$r = \frac{\mathbf{w}^T\mathbf{x}_0 + b}{\parallel \mathbf{w} \parallel} \ where\ r \lesseqgtr 0\ if\ \mathbf{w}^T\mathbf{x}_0 + b \lesseqgtr 0.$$

**Proof.**
Let $z\ s.t.\ \mathbf{w}^T\mathbf{z} + b = 0$ (i.e. a vector on the hyperplane.)

$$r = (\mathbf{x}_0 - \mathbf{z})^T \frac{\mathbf{w}}{\parallel \mathbf{w} \parallel} = \frac{\mathbf{w}^T\mathbf{x}_0 - \mathbf{w}^T\mathbf{z}}{\parallel \mathbf{w} \parallel} = \frac{\mathbf{w}^T\mathbf{x}_0 + b}{\parallel \mathbf{w} \parallel} \ ( \because\ \mathbf{w}^T\mathbf{z} = -b\ ) \ \blacksquare$$

#### Definition.A.1.16 Moore-penrose pseudoinverse matrix

Let $A$ be $m \times n$ matrix.

$$A^+ = \lim_{\alpha \rightarrow 0} (A^T A + \alpha I)^{-1} A^T$$

$A^+$ is called psedoinverse matrix of $A$.

**Theorem.A.1.4 Singular value decomposition with pseduinverse matrix**

Suppose $A$ be $m \times n$ matrix.

$$A^+ = VD^+U^T$$

Columns of $U_{m \times m}$ are left-singular vector of $A$. (Eigenvectors of $AA^T$)
Columns of $V_{n \times n}$ are right-singular vector of $A$. (Eigenvectors of $A^TA$)

**Proof.**

$$\begin{aligned} A^+ &= \lim_{\alpha \to 0}(A^TA + \alpha I)^{-1}A^T \\ &= \lim_{\alpha \to 0}(V\Sigma^TU^TU\Sigma V^T + \alpha I)^{-1}V\Sigma^TU^T \quad (\because \text{ SVD of } A) \\ &= \lim_{\alpha \to 0}(V\Sigma^T\Sigma V^T + \alpha I)^{-1}V\Sigma^TU^T \\ &= (V^T)^{-1}(\Sigma^T\Sigma)^{-1}V^{-1}V\Sigma^TU^T \\ &= V(\Sigma^T\Sigma)^{-1}\Sigma^TU^T \\ &= V\lim_{\alpha \to 0}(\Sigma^T\Sigma + \alpha I)^{-1}\Sigma^TU^T \\ &= V\Sigma^+U^T \qquad \blacksquare \end{aligned}$$

**Lemma.A.1.1 Woodbury formula**

Let $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{k \times k}$ are invertible and $U,\ V \in \mathbb{R}^{n \times k}$.
$A + UCV^T$ is invertible, if and only if, $C^{-1} + V^TA^{-1}U$ is invertible.
In addition,

$$(A + UCV^T)^{-1} = A^{-1} - A^{-1}U(C^{-1} + V^TA^{-1}U)^{-1}V^TA^{-1}$$

**Proof.**
$\det\begin{pmatrix} A & U \\ V^T & -C^{-1} \end{pmatrix} \neq 0 \quad (\because \ A \text{ and } -C^{-1})$
Therefore, there is an unique solution about following equation.

$$\begin{bmatrix} A & U \\ V^T & -C^{-1} \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} I \\ O \end{bmatrix} \quad \text{where } X \in \mathbb{R}^{n \times n},\ \ Y \in \mathbb{R}^{k \times n}$$

$$\begin{cases} AX + UY = I & \cdots (1) \\ V^TX - C^{-1}Y = O & \cdots (2) \end{cases}$$

$Y = CV^TX \quad (\because \ (2)) \quad \cdots (3)$
$AX + UCV^TX = (A + UCV^T)X = I \quad (\because \ (1),\ (3))$
$(A + UCV^T)^{-1} = X \quad (\because \ \text{definition of inverse matrix})$
$X = A^{-1}(I - UY) \quad (\because \ (1)) \quad \cdots (4)$
$V^TA^{-1}(I - UY) - C^{-1}Y = 0 \quad (\because \ (2),\ (4))$
$Y = (C^{-1} + V^TA^{-1}U)^{-1}V^TA^{-1} \quad \cdots (5)$

$$\begin{aligned} X &= A^{-1}(I - U(C^{-1} + V^TA^{-1}U)^{-1}V^TA^{-1}) \quad (\because \ (4),\ (5)) \\ &= A^{-1} - A^{-1}U(C^{-1} + V^TA^{-1}U)^{-1}V^TA^{-1} \qquad \blacksquare \end{aligned}$$

**Reference :**

Deep Learning - Yosha Benjio
대학수학2 - 박찬녕
https://proofwiki.org (https://proofwiki.org)
https://wikipedia.org (https://wikipedia.org)
https://ko.wikipedia.org/wiki/위키백과:TeX_문법
(https://ko.wikipedia.org/wiki/%EC%9C%84%ED%82%A4%EB%B0%B1%EA%B3%BC:TeX_%EB%AC%B8%EB%B2%95)