1 Basic Linear Algebra

- $\boldsymbol{A}^{-T} \triangleq \left(\boldsymbol{A}^T\right)^{-1} = \left(\boldsymbol{A}^{-1}\right)^T$
- $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$
- $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, iff $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible
- Frobenius norm $\|\boldsymbol{A}\|_F = \sqrt{\mathrm{Tr}\left(\boldsymbol{A}\boldsymbol{A}^T\right)}$
- **trace:**
  - $\mathrm{Tr}\left(\boldsymbol{A}^T\right) = \mathrm{Tr}\left(\boldsymbol{A}\right)$
  - $\mathrm{Tr}\left(\boldsymbol{A}+\boldsymbol{B}\right) = \mathrm{Tr}\left(\boldsymbol{A}\right) + \mathrm{Tr}\left(\boldsymbol{B}\right)$
  - $\mathrm{Tr}\left(\boldsymbol{A}\boldsymbol{B}^T\right) = \mathrm{Tr}\left(\boldsymbol{A}^T\boldsymbol{B}\right) = \sum_{i,j}\left(\boldsymbol{A}\circ\boldsymbol{B}\right)_{(i,j)}$
  - $\mathrm{Tr}\left(\boldsymbol{ABC}\right) = \mathrm{Tr}\left(\boldsymbol{CAB}\right) = \mathrm{Tr}\left(\boldsymbol{BCA}\right)$
- **determinant:**
  - $\det\boldsymbol{A}^T = \det\boldsymbol{A}$
  - $\det\boldsymbol{A}^{-1} = (\det\boldsymbol{A})^{-1}$
  - $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$, for square matrices of equal size.
  - If $\boldsymbol{A}$ is a triangular matrix (lower triangular or upper triangular),

$$\det\boldsymbol{A} = \prod_i A_{(i,i)}$$

2 Cholesky decomposition

The Cholesky decomposition of a positive-definite matrix A is a decomposition of the form:

$$\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^T$$

where $\boldsymbol{L}$ is a lower triangular matrix

- $\boldsymbol{\Sigma}^{-1} = \left(\boldsymbol{L}\boldsymbol{L}^T\right)^{-1} = \boldsymbol{L}^{-T}\boldsymbol{L}^{-1}$
- $\boldsymbol{L}\setminus\boldsymbol{x} \triangleq \boldsymbol{L}^{-1}\boldsymbol{x}$
- $\boldsymbol{\Sigma}^{-1}\boldsymbol{x} = \boldsymbol{L}^T\setminus\boldsymbol{L}\setminus\boldsymbol{x}$
- $\boldsymbol{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x} = \|\boldsymbol{L}\setminus\boldsymbol{x}\|_2^2$

- $\mathrm{Tr}\left(\boldsymbol{\Sigma}_b^{-1}\boldsymbol{\Sigma}_a\right) = \|\boldsymbol{L}_b \setminus \boldsymbol{L}_a\|_F^2$
  Proof:

$$\begin{aligned}\mathrm{Tr}\left(\boldsymbol{\Sigma}_b^{-1}\boldsymbol{\Sigma}_a\right) &= \mathrm{Tr}\left(\left(\boldsymbol{L}_b\boldsymbol{L}_b^T\right)^{-1}\boldsymbol{L}_a\boldsymbol{L}_a^T\right)\\ &= \mathrm{Tr}\left(\boldsymbol{L}_b^{-T}\boldsymbol{L}_b^{-1}\boldsymbol{L}_a\boldsymbol{L}_a^T\right) = \mathrm{Tr}\left(\boldsymbol{L}_a^T\boldsymbol{L}_b^{-T}\boldsymbol{L}_b^{-1}\boldsymbol{L}_a\right)\\ &= \mathrm{Tr}\left(\left(\boldsymbol{L}_b^{-1}\boldsymbol{L}_a\right)^T\left(\boldsymbol{L}_b^{-1}\boldsymbol{L}_a\right)\right) = \|\boldsymbol{L}_b \setminus \boldsymbol{L}_a\|_F^2\end{aligned}$$

- $\log\left(|\boldsymbol{\Sigma}|\right) = 2\sum_i \log\left(L_{(i,i)}\right) = 2\mathrm{Tr}\left(\log\left(\boldsymbol{L}\right)\right)$

3 Inverse

$$(\boldsymbol{I}+\boldsymbol{P})^{-1} = \boldsymbol{I} - (\boldsymbol{I}+\boldsymbol{P})^{-1}\boldsymbol{P} \tag{1}$$

$$(\boldsymbol{I}+\boldsymbol{P}\boldsymbol{Q})^{-1}\boldsymbol{P} = \boldsymbol{P}(\boldsymbol{I}+\boldsymbol{Q}\boldsymbol{P})^{-1} \tag{2}$$

3.1 Matrix inversion lemma (Sherman-Morrison-Woodbury)

- $$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{D}\boldsymbol{A}^{-1} \tag{3}$$
- $(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{I}+\boldsymbol{C}\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{C}\boldsymbol{D}\boldsymbol{A}^{-1}$
- $\left(\boldsymbol{A}+\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^T\right)^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{X}\left(\boldsymbol{B}^{-1}+\boldsymbol{X}^T\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{X}^T\boldsymbol{A}^{-1}$
- $(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}\boldsymbol{B}\boldsymbol{C} = \boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}$, using 3 and 1 2
- $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1}\left(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}\right)^{-1}\boldsymbol{B}^{-1}$
- $\left(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}\right)^{-1} = \boldsymbol{A}\left(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}\right)\boldsymbol{B}$
  where $\boldsymbol{A}$ and $\boldsymbol{B}$ are square and invertible matrices.

4 square

$$\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x} - 2\boldsymbol{b}^T\boldsymbol{x} = \left(\boldsymbol{x}-\boldsymbol{M}^{-1}\boldsymbol{b}\right)^T\boldsymbol{M}\left(\boldsymbol{x}-\boldsymbol{M}^{-1}\boldsymbol{b}\right) - \boldsymbol{b}^T\boldsymbol{M}^{-1}\boldsymbol{b} \tag{4}$$

A Notation

Notation	Description
$\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$	Tensor of order N
$x, \boldsymbol{x}, \boldsymbol{X}$	Scalar, vector and matrix. Non-bold letters do not strictly represent scalars, in many cases their meaning should be extracted from context
$x_{i_1,i_2,...,i_N}$, $\underline{\mathbf{X}}_{(i_1,i_2,...,i_N)}$	$(i_1, i_2, ..., i_N)$th entry of $\underline{\mathbf{X}}$
$\boldsymbol{x}_{:,i}$, $\boldsymbol{X}_{(:,i)}$	ith column of the matrix $\boldsymbol{X}$. Colons are used for indexing an entire dimension.
$\boldsymbol{x}_{:,i_2,...,i_N}$, $\underline{\mathbf{X}}_{(:,i_2,...,i_N)}$	Mode-1 fiver of $\underline{\mathbf{X}}$
$\boldsymbol{X}_{:,:,...,i_N}$, $\underline{\mathbf{X}}_{(:,:,...,i_N)}$	Frontal slice of $\underline{\mathbf{X}}$
$\underline{\mathbf{X}}_{(+,:,...,:)}$	Partial sum-reduction over first dimension
$\underline{\mathbf{X}}_{(+)}$	Sum-reduction over all elements in the tensor
$\underline{\mathbf{1}}$	Tensor whose elements are equal to one. Their order and size is usually extracted from context
$\underline{\mathbf{A}} \odot \underline{\mathbf{B}}$	Element-wise product between tensors $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$
$\underline{\mathbf{C}}_{(i,j)} = \underline{\mathbf{A}}_{(i,k)}\underline{\mathbf{B}}_{(k,j)}$	Tensor contraction using Einstein notation. In this case is just the matrix multiplication between $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$

Table 1: Notation for vectors, matrices and tensors