Linear algebra review

- Basic notation and review
- Common operators and decompositions

Notation

- We denote *scalars* by an italicized variable: e.g., a or A.
- We denote *vectors* by a lowercase bolded variable: e.g., $\mathbf{a}$.
- We denote *matrices* by an uppercase bolded variable: e.g., $\mathbf{A}$.
- We denote *tensors* by an uppercase sans serif variable: e.g., A.

Vectors

Let

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

be a column vector with n elements. A column vector is distinct from a row vector with n elements, which we denote: $\mathbf{z} = [z_1 \ z_2 \ \ldots \ z_n]$.

- The transpose of a column vector is a row vector, and vice versa. Concretely, $\mathbf{x} = [x_1 \ x_2 \ \ldots \ x_n]^T$.
- The dot product of two column vectors, $\mathbf{x}$ and $\mathbf{y}$, is given by

$$\mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

- $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal* if $\mathbf{x}^T\mathbf{y} = 0$.

Vector norm

- The norm of a vector measures its length. The *p-norm* of a vector is given by:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

  for $p \geq 1$. We will almost always care about the *Euclidean* norm, which is the 2-norm.

- The *Euclidean* norm can also be written as:

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T \mathbf{x}}$$

- When we write $\|\mathbf{x}\|$ without a subscript, it can be assumed to be the Euclidean norm.
- The dot product can be written as:

$$\mathbf{x}^T \mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos \theta$$

  where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

Matrices

Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

be an $m \times n$ matrix. Note that we will sometimes denote the i, j element of $\mathbf{A}$ as $\mathbf{A}_{ij}$. For the above matrix, $\mathbf{A}_{ij} = a_{ij}$.

- The transpose of $\mathbf{A}$, denoted $\mathbf{A}^T$, satisfies:

$$\mathbf{A}_{ij} = (\mathbf{A}^T)_{ji}$$

- If the matrix is square, i.e., $m = n$, and has rank n, then the inverse of $\mathbf{A}$, denoted $\mathbf{A}^{-1}$, satisfies

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

  where $\mathbf{I}$ is an $n \times n$ identity matrix.
- A matrix is *symmetric* if $\mathbf{A} = \mathbf{A}^T$.

Matrix trace

- The trace of a matrix, denoted $\text{tr}(\mathbf{A})$, is the sum of its diagonal elements.

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$$

- The trace operator is invariant to transposition.

$$\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$$

- The trace operator is invariant to cyclic permutations of its input. For example,

$$\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CAB}) = \text{tr}(\mathbf{BCA})$$

- The trace operator is linear.

$$\text{tr}(a\mathbf{X} + b\mathbf{Y}) = a\text{tr}(\mathbf{X}) + b\text{tr}(\mathbf{Y})$$

Eigendecomposition

- An eigenvector, $\mathbf{u}_i$, and its corresponding eigenvalue, λ_i, of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ satisfy:
$$\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i$$
- If we collect all of $\mathbf{A}$'s eigenvectors and eigenvalues into the following matrices,
$$\mathbf{U} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & \cdots & | \end{bmatrix} \qquad \text{and} \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$
we have that $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$. This is the eigendecomposition of $\mathbf{A}$.
- If $\mathbf{A}$ is *normal*, then its eigenvectors are *orthonormal*, i.e., $\mathbf{u}_i^T\mathbf{u}_j = 0$ for $i \neq j$ and $\mathbf{u}_i^T\mathbf{u}_i = 1$.

Singular value decomposition

- Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the singular value decomposition of $\mathbf{A}$ is:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$$

  where $\mathbf{U}$ is an $m \times m$ matrix with orthonormal columns and $\mathbf{V}$ is an $n \times n$ matrix with orthonormal columns. The matrix Σ is $m \times n$ with σ_i as its ith diagonal element.
- The columns of $\mathbf{U}$ are called the left singular vectors of $\mathbf{A}$.
- The columns of $\mathbf{V}$ are called the right singular vectors of $\mathbf{A}$.
- σ_i is called the ith singular value of $\mathbf{A}$.

Other matrix properties

- Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The Frobenius norm of $\mathbf{A}$ is

$$\begin{aligned} \|\mathbf{A}\|_F &= \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2} \\ &= \sqrt{\operatorname{tr}(\mathbf{A}\mathbf{A}^T)} \end{aligned}$$

- Consider a symmetric matrix, $\mathbf{A}$.
  - $\mathbf{A}$ is called positive definite if $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ for all $\mathbf{x}$.
  - If $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$, $\mathbf{A}$ is positive semidefinite.
  - Analogous definitions exist for negative definite and negative semidefinite matrices.