CS205 Review Session #2 Notes

Symmetric matrices and dot product

Lemma An $n \times n$ matrix $\mathbf{A}$ is symmetric $\iff \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n\ \mathbf{x} \cdot \mathbf{A}\mathbf{y} = \mathbf{y} \cdot \mathbf{A}\mathbf{x}$

Proof In the forward direction, if $\mathbf{A}$ is symmetric we have:

$$\mathbf{x} \cdot \mathbf{A}\mathbf{y} = \mathbf{x}^T\mathbf{A}\mathbf{y} = (\mathbf{x}^T\mathbf{A}\mathbf{y})^T = \mathbf{y}^T\mathbf{A}^T\mathbf{x} = \mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{y} \cdot \mathbf{A}\mathbf{x}$$

For the converse, recall that $\mathbf{e}_i$ is the i-th cartesian basis vector. Taking $\mathbf{x} = \mathbf{e}_i$ and $\mathbf{y} = \mathbf{e}_j$ and then the reverse, we have:

$$\begin{aligned} \mathbf{e}_i \cdot A\mathbf{e}_j = \mathbf{e}_i^T\mathbf{A}\mathbf{e}_j = A_{ij} \\ \mathbf{e}_j \cdot A\mathbf{e}_i = \mathbf{e}_j^T\mathbf{A}\mathbf{e}_i = A_{ji} \end{aligned}$$

Thus $A_{ij} = A_{ji}$ for all i, j and $\mathbf{A}$ is symmetric.

Lemma If $\mathbf{A}, \mathbf{B}$ are symmetric $n \times n$ matrices and $\forall \mathbf{x} \in \mathbb{R}^n\ \mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{B}\mathbf{x}$ then $\mathbf{A} = \mathbf{B}$

Proof First, take $\mathbf{x} = \mathbf{e}_i$:

$$\mathbf{e}_i^T\mathbf{A}\mathbf{e}_i = \mathbf{e}_i^T\mathbf{B}\mathbf{e}_i \Rightarrow A_{ii} = B_{ii}$$

Similarly, taking $\mathbf{x} = \mathbf{e}_i + \mathbf{e}_j$ gives:

$$\begin{aligned} (\mathbf{e}_i + \mathbf{e}_j)^T A(\mathbf{e}_i + \mathbf{e}_j) &= (\mathbf{e}_i + \mathbf{e}_j)^T B(\mathbf{e}_i + \mathbf{e}_j) \\ A_{ii} + A_{jj} + A_{ij} + A_{ji} &= B_{ii} + B_{jj} + B_{ij} + B_{ji} \\ 2A_{ij} &= 2B_{ij} \end{aligned}$$

Fundamental Subspaces

Recall that an $n \times n$ matrix $\mathbf{A}$ is just a basis for a vector space. If the columns of $\mathbf{A}$ are all linearly independent, then $\mathbf{A}$ is a basis for $\mathbb{R}^n$. If some columns are linear combinations of others, then the maximal subset of columns that are linearly independent forms a basis for a subspace. The subspace spanned by the columns is called the **column space** of the matrix, and is given as $\text{col}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$. The **rank** of $\mathbf{A}$ is the number of linearly independent columns, and when $\mathbf{A}$ has full rank $\text{col}(\mathbf{A}) = \mathbb{R}^n$.

When $\mathbf{A}$ is rank deficient, it is said to have a non-trivial **null space**. The null space of a matrix $\mathbf{A}$ is defined to be $\text{NS}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$. If $\text{NS}(\mathbf{A}) = \{\mathbf{0}\}$ then, we say that $\mathbf{A}$ has a trivial null space. In general, $\mathbf{A}$ has a trivial null space if $\mathbf{A}$ is square and non-singular. The minimum number of columns that must be removed from a matrix to make the remaining columns linearly independent is called the **nullity** or **kernel** of the matrix.

Now let us consider the more general case where $\mathbf{A}$ is $m \times n$. This allows us to more clearly understand the dual nature of the column space (given by $\text{col}(\mathbf{A}) = \{\mathbf{Ax} \in \mathbb{R}^m \mid \mathbf{x} \in \mathbb{R}^n\}$) and the null space (given by $\text{NS}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{0}\}$) of $\mathbf{A}$ by examining their respective dimensions. In particular, the **Rank-nullity theorem** states that $\text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n$. The definitions for the column space and null space of $\mathbf{A}^T$ follow directly.

The net result of this analysis leads us to conclude that there are four fundamental subspaces. The first and the second are the column space and null space defined above. The next is the **row space**, which is the column space of $\mathbf{A}^T$ or the space spanned the rows of $\mathbf{A}$. The last fundamental subspace is the **left null space** of $\mathbf{A}$ which is the null space of $\mathbf{A}^T$. The null space of $\mathbf{A}$ and row space of $\mathbf{A}$ are subspaces of $\mathbb{R}^n$, while the left null space of $\mathbf{A}$ and column space of $\mathbf{A}$ are subspaces of $\mathbb{R}^m$.

The **Fundamental Theorem of Linear Algebra** states that the column space of $\mathbf{A}$ and the null space of $\mathbf{A}^T$ are **orthogonal complements** of one another:

$$\text{col}(\mathbf{A}) = \text{NS}(\mathbf{A}^T)^\perp$$

This means that $\forall \mathbf{x} \in \text{col}(\mathbf{A})\ \forall \mathbf{y} \in \text{NS}(\mathbf{A}^T)\ \mathbf{x}^T\mathbf{y} = 0$. Additionally, the dimension of $\mathbf{A}$'s column space will be $\text{rank}(\mathbf{A})$ and the dimensional of $\mathbf{A}^T$'s null space will be $m - \text{rank}(\mathbf{A})$. Similarly, the *row space* of $\mathbf{A}$ is the orthogonal complement of the *null space* of $\mathbf{A}$. Of course, if $\mathbf{A} = \mathbf{A}^T$, the row space and the column space are the same.

Why is this important? Given an $m \times n$ matrix $\mathbf{A}$, we can use this analysis to decompose an arbitrary vector $\mathbf{x} \in \mathbb{R}^m$ into $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ where $\mathbf{x}_1 = \mathbf{Ay}$ (for some $\mathbf{y} \in \mathbb{R}^n$) and $\mathbf{A}^T\mathbf{x_2} = 0$. Similarly, we can decompose an arbitrary $\mathbf{z} \in \mathbb{R}^n$ into $\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2$ with $\mathbf{z}_1 = \mathbf{A}^T\mathbf{w}$ (for some $\mathbf{w} \in \mathbb{R}^m$) and $\mathbf{Az}_2 = 0$. Moreover we know that $\mathbf{x}_1^T\mathbf{x}_2 = 0$ and $\mathbf{z}_1^T\mathbf{z}_2 = 0$ because these spaces are complimentary.

Projection and Reflection

Consider a vector subspace V of $\mathbb{R}^n$. Its normal complement is the vector space $V^\perp$, which contains all vectors of $\mathbb{R}^n$ that are normal to some vector of V. We know that every vector $\mathbf{x} \in \mathbb{R}^n$ can be decomposed as:

$$\mathbf{x} = \mathbf{x_1} + \mathbf{x_2},\ \mathbf{x_1} \in V,\ \mathbf{x_2} \in V^\perp$$

The *projection* matrix $\mathbf{P}$ onto the subspace $\mathbf{V}$ operates on $\mathbf{x}$ as follows:

$$\mathbf{Px} = \mathbf{P}(\mathbf{x_1} + \mathbf{x_2}) = \mathbf{Px_1} + \mathbf{Px_2} = \mathbf{x_1} + \mathbf{0} = \mathbf{x_1}$$

The *reflection* matrix $\mathbf{R}$ with respect to the subspace $\mathbf{V}$ has the following effect on $\mathbf{x}$:

$$\mathbf{Rx} = \mathbf{R}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{Rx}_1 + \mathbf{Rx}_2 = \mathbf{x}_1 + (-\mathbf{x}_2) = \mathbf{x}_1 - \mathbf{x}_2$$

Diagonally Dominant Matrices

We have seen that diagonally dominant matrices are a class of generally well-conditioned matrices. In particular they do not require pivoting during Gaussian elimination in order to enforce stability. Note, however, that our definition of diagonal dominance may differ slightly from the literature.

As an exercise, we will prove that 2×2 diagonally dominant matrices with positive diagonal elements are positive definite. Consider the matrix:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

with $a_{11}, a_{22} > 0$. Diagonal dominance for this matrix translates to

$$\begin{aligned} a_{11} = |a_{11}| > |a_{12}|, |a_{21}| \\ a_{22} = |a_{22}| > |a_{12}|, |a_{21}| \end{aligned}$$

If $\mathbf{x} = (x_1, x_2)^T$, we have:

$$\begin{aligned} \mathbf{x^T A x} &= a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_1x_2 + a_{22}x_2^2 \\ &\geq |a_{11}||x_1|^2 + |a_{22}||x_2|^2 - (|a_{12}| + |a_{21}|)|x_1||x_2| \end{aligned}$$

This implies that:

$$\begin{aligned} 2\mathbf{x^T A x} &\geq 2|a_{11}||x_1|^2 + 2|a_{22}||x_2|^2 - 2(|a_{12}| + |a_{21}|)|x_1||x_2| \\ &> (|a_{12}| + |a_{21}|)|x_1|^2 + (|a_{12}| + |a_{21}|)|x_2|^2 - (2|a_{12}| + 2|a_{21}|)|x_1||x_2| \\ &= |a_{12}|(|x_1| - |x_2|)^2 + |a_{21}|(|x_1| - |x_2|)^2 \geq 0 \end{aligned}$$

Which is precisely the criteria positive definiteness.