Lecture 8: Solutions to $Ax = y$ and Normed Linear Spaces

2.7.3 Finding $\mathcal{N}\{A\}$

To find $\mathcal{N}\{A\}$, we need to characterize all solutions to $Ax = 0$. Recall that row operations preserve $\mathcal{N}\{A\}$, so that $Ax = 0 \iff \tilde{Ax} = 0$. We can solve $\tilde{Ax} = 0$ recursively backwards starting with the last row. The solution can be simplified by realizing that the components of $x$ corresponding to the columns of $\tilde{A}$ that do not have a leading entry can be selected arbitrarily.

Example:

Let $\tilde{A} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 \end{bmatrix}$ be the echelon form of some matrix $A$. We want to find $\mathcal{N}\{A\} = \mathcal{N}\{\tilde{A}\}$.

\[
\begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{bmatrix} = 0
\]

We have $\text{rank}\{A\} = 2$, $\text{nullity}\{A\} = 3$. The components of $\xi$ corresponding to the columns with indices in $K^c$ are $\xi_3$, $\xi_4$, $\xi_5$, and can be selected arbitrarily. To find three linearly independent solutions that span $\mathcal{N}\{A\}$, we set $\xi_3$, $\xi_4$, $\xi_5$ to the canonical basis vectors of $\mathbb{R}^3$, and solve for the remaining two components:

\[
x^1 = \begin{bmatrix} \xi_1^1 \\ \xi_2^1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad x^2 = \begin{bmatrix} \xi_1^2 \\ \xi_2^2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad x^3 = \begin{bmatrix} \xi_1^3 \\ \xi_2^3 \\ 0 \\ 0 \\ 1 \end{bmatrix}
\]

1) $\tilde{Ax}^1 = 0 : \xi_1^1 + \xi_2^1 = 0, \xi_2^1 = 0$
   $\Rightarrow \xi_1^1 = 0, \xi_2^1 = 0$

2) $\tilde{Ax}^2 = 0 : \xi_1^1 + \xi_2^1 = -1, \xi_2^1 = -2$
   $\Rightarrow \xi_1^1 = 1, \xi_2^1 = -2$
\[ \tilde{A}x^3 = 0 : \quad \xi_1^1 + \xi_2^1 = -1, \xi_2^1 = -1 \]
\[ \Rightarrow \xi_1^1 = 0, \xi_2^1 = -1 \]

Therefore:

\[
\mathcal{N} \{ \mathcal{A} \} = \text{span} \begin{bmatrix}
0 \\
0 \\
1 \\
0 \\
0
\end{bmatrix}, \begin{bmatrix}
1 \\
0 \\
0 \\
0
\end{bmatrix}, \begin{bmatrix}
0 \\
0 \\
-1 \\
0
\end{bmatrix}
\]

### 2.7.4 The Non-Homogeneous Equation \( Ax = y \)

Here \( A, y \) are given and \( x \) is the unknown. Since

\[ Ax = y \Leftrightarrow Ax - y = 0 , \]

we have the equivalent homogeneous equation

\[
\begin{bmatrix}
A \\
-1
\end{bmatrix} \begin{bmatrix}
x \\
0
\end{bmatrix} = 0 ,
\]

or:

\[
W \begin{bmatrix}
x \\
-1
\end{bmatrix} = 0 .
\]

Thus, if \( A \in \mathcal{F}^{m \times n} \), then the augmented matrix \( W \) is \( m \times (n + 1) \).

The solutions \( \begin{bmatrix}
x \\
-1
\end{bmatrix} \) lie in \( \mathcal{N} \{ W \} \), but their last component is constrained to be -1. Therefore, they lie in a subset of \( \mathcal{N} \{ W \} \) which is a coset, e.g., a subspace plus a constant vector. To characterize the coset, let \( A_r \) denote the restriction of \( A \) to one of its support spaces \( S_{sp} (A) \) , say, \( \mathcal{D} \{ A \} / \mathcal{N} \{ A \} \), and to its range. Then there is a solution of \( Ax = y \) for \( x \in \mathcal{D} \{ A \} / \mathcal{N} \{ A \} \) given by:

\[ x = A_r^{-1} y . \]

Next, every solution of \( Ax = y \) must lie in the coset:

\[ A_r^{-1} y + \mathcal{N} \{ A \} \]
Proposition:

There exists (one or more) solutions of $Ax = y$

$$ (a) \iff y \in \mathcal{R}\{A\} $$

$$ (b) \iff \text{rank}\{A\} = \text{rank}\{W\} $$

Proof:

(a) This is clear.

(b) $\exists x : y = Ax \iff y$ is linearly dependent on columns of $A$

$$ \iff \text{number of independent columns of } A = \text{number of independent columns of } [A \ y] $$

$$ \iff \text{rank}\{A\} = \text{rank}\{[A \ y]\} = \text{rank}\{W\} $$

Suppose $x$ is a solution of $Ax = y$. Choose any support spaces $S_{sp}(A)$, say, $\mathcal{D}\{A\} / \mathcal{N}\{A\}$

Decompose the vector as:

$$ x = x_r + x_n, \quad x_r \in S_{sp}\{A\}, x_n \in \mathcal{N}\{A\} $$
\[ Ax = Ax_r + Ax_n = A_r x_r = y \]

Since \( A_r \) is invertible, we have \( x_r = A_r^{-1} y \), and:

\[ x = A_r^{-1} y + x_n, \]

If \( x_1, x_2 \) are two solutions:

\[ Ax_1 = y, \quad Ax_2 = y \]

\[ A(x_1 - x_2) = 0 \implies x_1 - x_2 \in \mathcal{N}\{A\} \]

Hence, every solution is of the form:

\[ A_r^{-1} y + \mathcal{N}\{A\} \]

Conclusion:

Every solution of \( Ax = y \) lies in the coset

\[ \left[ A_r^{-1} y + \mathcal{N}\{A\} \right] \]

Proposition:

The number of independent solutions of \( Ax = y \) is given by:

\[ \dim \{\mathcal{N}\{A\}\} = n - \text{rank}\{A\} = n - \text{rank}\{W\} \].
3 Normed Linear Spaces

The norm on a linear space is a natural generalization of the length of a vector on the Euclidean space $E^n$ over the field $\mathbb{R}$, e.g., the linear space $(\mathbb{R}^n, \mathbb{R})$ on which $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$.

3.1 Inner-Product, Norm, Metric Spaces

3.1.1 Definition: Inner Product Space

An inner product space is a vector $(\mathcal{V}, \mathcal{F})$ on which is defined a function

$$\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathcal{F}$$

such that the following axioms are satisfied $\forall x, y, z \in \mathcal{V}, \alpha \in \mathcal{F}$:

(i) $\langle x, y \rangle = \langle y, x \rangle$ if $\mathcal{F} = \mathbb{R}$ or $\langle x, y \rangle = \overline{\langle y, x \rangle}$ if $\mathcal{F} = \mathbb{C}$.

(ii) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$

(iii) $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$

(iv) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \iff x = \theta$

The norm of $x \in \mathcal{V}$ associated with an inner product space is given by: $\|x\| := \sqrt{\langle x, x \rangle}$.

3.1.2 Definition: Normed Linear Space

A normed linear space is an ordered pair $(\mathcal{V}, \|\cdot\|)$ where $(\mathcal{V}, \mathcal{F})$ is a vector space on which is defined a function

$$\|\cdot\| : \mathcal{V} \to \mathcal{F}$$

such that the following axioms are satisfied $\forall x, y, z \in \mathcal{V}, \alpha \in \mathcal{F}$:

(i) $\|x\| \geq 0; \|x\| = 0 \iff x = \theta$

(ii) $\|\alpha x\| = |\alpha| \|x\|$ (sublinearity)

(iii) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)
If we consider for instance \( V = \mathbb{R}^n, F = \mathbb{R} \) on which \( \|x\| := \left( \sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} \), then:

(i) states that only the zero vector has zero length and every other vector has positive length,

(ii) states that if a vector is multiplied by a real scalar, then the length of the vector gets multiplied by the absolute value of the scalar,

(iii) states that the length of the sum of two vectors is no larger than the sum of their lengths, e.g.,

\[
\|x + y\| \leq \|x\| + \|y\|.
\]

**Examples:**

(a) \((\mathbb{R}^n, \|\cdot\|_{\infty})\), \( \|x\|_{\infty} := \max_{1 \leq i \leq n} |x_i| \) which is the \( l_{\infty} \)-norm on \( \mathbb{R}^n \).

Check (iii): Let \( x = [x_1 \cdots x_n]^T, y = [y_1 \cdots y_n]^T \in \mathbb{R}^n \). Then,

\[
|x_i + y_i| \leq |x_i| + |y_i|, \quad i = 1, \ldots, n \quad \text{(triangle inequality of real numbers)}
\]

\[
\Rightarrow \quad \|x + y\|_{\infty} = \max_{1 \leq i \leq n} |x_i + y_i| \leq \max_{1 \leq i \leq n} (|x_i| + |y_i|)
\]

\[
\leq \max_{1 \leq i \leq n} |x_i| + \max_{1 \leq i \leq n} |y_i| = \|x\|_{\infty} + \|y\|_{\infty}
\]

(b) \((\mathbb{R}^n, \|\cdot\|_1)\), \( \|x\|_1 := \sum_{i=1}^{n} |x_i| \) which is the \( l_1 \)-norm on \( \mathbb{R}^n \).

(c) \((\mathbb{R}^n, \|\cdot\|_p)\), \( \|x\|_p := \sum_{i=1}^{n} |x_i|^p \) which is the \( l_p \)-norm on \( \mathbb{R}^n \), \( p \in [1, +\infty] \).

The \( l_p \)-norms typically used are the \( l_1 \)-norm, the \( l_2 \)-norm, and the \( l_{\infty} \)-norm.

(d) \((\mathbb{C}^n, \|\cdot\|_p)\), \( \|x\|_p := \sum_{i=1}^{n} |x_i|^p \) which is the \( l_p \)-norm on \( \mathbb{C}^n \), \( p \in [1, +\infty] \).
Notes:

Both $\mathbb{R}^n$ and $\mathbb{C}^n$ are examples of finite-dimensional linear vector spaces, and consequently it can be shown that given any two norms $\|x\|_a, \|x\|_b$ on $\mathbb{R}^n$ there exist constants $k_1, k_2$ such that:

$$k_1 \|x\|_a \leq \|x\|_b \leq k_2 \|x\|_a, \quad \forall x \in \mathbb{R}^n \ (\text{or } \mathbb{C}^n)$$

for example:

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty, \quad \forall x \in \mathbb{R}^n,$$

and,

$$\|x\|_\infty \leq \|x\|_\infty \leq n \|x\|_\infty, \quad \forall x \in \mathbb{R}^n.$$

3.1.3 Metric on $\mathcal{V}$ and Convergence

We can now use the inner product to define a topology, e.g., a notion of convergence in $\mathcal{V}$. Thus, we can think of the quantity $\|x - y\|$ as the distance between vectors $x$ and $y$, defined by:

$$d(\bullet,\bullet) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}_+$$

$$d(x, y) := \|x - y\|, \quad x, y \in \mathcal{V}$$

This is called a metric on space $\mathcal{V}$. The properties of this metric follow from properties of the inner product and the triangle inequality on $\mathcal{V}$:

(i) \quad $d(x, y) = d(y, x)$,

(ii) \quad $d(x, y) = 0 \iff y = x$,

(iii) \quad $d(x, z) \leq d(x, y) + d(y, z)$.

Definition: Convergence in $(\mathcal{V}, \|\|)$

A sequence of vector $\{x_i\}_{i=1}^\infty$ in $(\mathcal{V}, \|\|)$ is said to converge to $x_0 \in \mathcal{V}$ if for every $\varepsilon > 0$, $\exists$ an integer $N(\varepsilon)$ such that:

$$d(x_i, x_0) = \|x_i - x_0\| < \varepsilon, \quad \forall i \geq N(\varepsilon).$$

That is, the distance between the $i^{th}$ vector and its limit tends to zero as $i$ tends to infinity.
Note:

In order to test the convergence of a given sequence of vectors, we need to know its limit, which is often difficult to have (and is kind of a chicken and egg problem.) Therefore, we need a test for convergence without the knowledge of \( x_0 \). This is provided by the concept of a Cauchy sequence.

Definition: Cauchy Sequence

A sequence \( \{ x_i \}_{i=1}^{\infty} \) in \( (V, \|\cdot\|) \) is said to be a Cauchy sequence if for every \( \varepsilon > 0 \), \( \exists \) an integer \( N(\varepsilon) \) such that:

\[
d(x_k, x_n) = \| x_k - x_n \| < \varepsilon, \quad \forall k, n \geq N(\varepsilon).
\]

Thus, a sequence is convergent if its terms approach arbitrarily closely a fixed element, whereas a sequence is Cauchy if its terms approach each other arbitrarily closely.

Lemma:

Every convergent sequence in \( (V, \|\cdot\|) \) is a Cauchy sequence.

Proof:

Let \( \{ x_i \}_{i=1}^{\infty} \) be a convergent sequence in \( (V, \|\cdot\|) \) with limit \( x_0 \). To show that the sequence is also Cauchy, suppose \( \varepsilon > 0 \) is given, and pick an integer \( N(\varepsilon) \) such that \( \| x_i - x_0 \| < \frac{\varepsilon}{2} \), \( \forall i \geq N(\varepsilon) \).

Then, by the triangle inequality,

\[
\| x_i - x_j \| \leq \| x_i - x_0 \| + \| x_0 - x_j \| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}, \quad \forall i, j \geq N(\varepsilon).
\]

Hence, sequence \( \{ x_i \}_{i=1}^{\infty} \) is Cauchy.

The above lemma shows that if the elements of a sequence \( \{ x_i \}_{i=1}^{\infty} \) in \( (V, \|\cdot\|) \) are getting closer and closer to a fixed element, then they must at the same time be getting closer and closer to each other.

The next question addresses the convergence of a Cauchy sequence.

Question: If the vectors are getting closer and closer to each other, are they also all getting closer to a fixed vector in \( (V, \|\cdot\|) \)?

Answer: Not in general, unfortunately.

Here is a counterexample to Cauchy sequence \( \Rightarrow \) convergent sequence.
Consider \((\mathcal{V}, \|\cdot\|)\) with \(\mathcal{V} := \text{space of all polynomials defined on } [0, 1] \text{ with real coefficients, and} \|x\| := \max_{t \in [0,1]} |x(t)|.\)

Let \(x_i : [0, 1] \to \mathcal{V}, \quad x_i(t) := \sum_{k=0}^{\infty} \frac{t^k}{k!}\) be the \(i\)th vector in the sequence \(\{x_i\}_{i=0}^{\infty}\).

(as an exercise, show that this sequence is Cauchy.)

We can see that there exists a function which is, in the norm \(\|\cdot\|\), the limit of this sequence, namely the exponential \(e^t\), however \(e^t \not\in \mathcal{V}\) as it is not a polynomial. That is, the sequence converges to a limit outside of the space.