Vector spaces

Paul Schrimpf

UBC
Economics 526

September 23, 2013

Overview

- Main idea: take intuition from $\mathbb{R}^3$ and apply to $\mathbb{R}^n$ and other vector spaces

Section 1

Normed vector spaces

Normed vector spaces

- Measure length or distance

Definition

A **normed vector space**, $(V, \mathbb{F}, +, \cdot, \|\cdot\|)$, is a vector space with a function, called the **norm**, from V to $\mathbb{F}$ and denoted by $\|v\|$ with the following properties:

1. $\|v\| \geq 0$ and $\|v\| = 0$ iff $v = 0$,
2. $\|\alpha v\| = |\alpha| \, \|v\|$ for all $\alpha \in \mathbb{F}$,
3. The triangle inequality holds:
$$\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$$
for all $v_1, v_2 \in V$.

- Shortest distance between two points is a straight line

Examples

Example

$\mathbb{R}^3$ is a normed vector space with norm

$$\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

This norm is exactly how we usually measure distance. For this reason, it is called the Euclidean norm.
More generally, for any n, $\mathbb{R}^n$, is a normed vector space with norm

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Examples

Example

$\mathbb{R}^n$ with the norm

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

for $p \in [1, \infty]$[1] is a normed vector space. This norm is called the p-norm.

[1] Where $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$

Examples

Example

$\mathcal{L}^p(0,1)$ with p-norm

$$\|f\|_p = \left(\int_0^1 |f(x)|^p dx \right)^{1/p}$$

is a normed vector space. Moreover, $\mathcal{L}^p(0,1)$ is a different space for different p. For example, $\frac{1}{x^{1/p}} \notin \mathcal{L}^p(0,1)$, but $\frac{1}{x^{1/p}} \in \mathcal{L}^q(0,1)$ for $q < p$.

Section 2

Inner product spaces

Inner product spaces

- Measure angles

Definition

A real **inner product space** is a vector space over the field $\mathbb{R}$ with an additional operation called the inner product that is function from $V \times V$ to $\mathbb{R}$. We denote the inner product of $v_1, v_2 \in V$ by $\langle v_1, v_2 \rangle$. It has the following properties:

1. Symmetry: $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$
2. Linear: $\langle av_1 + bv_2, v_3 \rangle = a \langle v_1, v_3 \rangle + b \langle v_2, v_3 \rangle$ for $a, b \in \mathbb{R}$
3. Positive definite: $\langle v, v \rangle \geq 0$ and equals 0 iff $v = 0$.

- Inner product space is also a normed vector space

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Example

- $\mathbb{R}^n$ with the dot product,

$$\langle x, y \rangle = x \cdot y = \sum_{i=1}^{n} x_i y_i = x^T y$$

- Norm induced by the inner product is the Euclidean norm

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Measuring angles

- $$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle\end{aligned}$$

- In $\mathbb{R}^n$ with the Euclidean norm when x and y are at right angles to one other, $\langle x,y\rangle = 0$, and we have the Pythagorean theorem:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 .$$

Definition

Let $x, y \in V$, an inner product space. x and y are **orthogonal** iff $\langle x,y\rangle = 0$.

Measuring angles

Theorem

Let $u, v \in \mathbb{R}^n$, then the angle between them is

$$\theta = \cos^{-1} \frac{\langle u, v \rangle}{\|u\| \, \|v\|}.$$

Proof.

Draw picture.

$$\cos\theta = \frac{\|tv\|}{\|u\|}.$$

Use Pythagorean theorem,

$$\begin{aligned} \|u\|^2 &= \|tv\|^2 + \|u - tv\|^2 \\ \|u\|^2 &= t^2\|v\|^2 + \|u\|^2 - 2t\langle u, v\rangle + t^2\|v\|^2 \\ 2t\langle u, v\rangle &= 2t^2\|v\|^2 \end{aligned}$$

Useful inequalities

- Triangle inequality

Theorem (Reverse triangle inequality)

Let V be a normed vector space and $x, y \in V$. Then

$$\left|\|x\| - \|y\|\right| \leq \|x - y\| .$$

Proof.

By the usual triangle inequality,

$$\begin{aligned} \|x\| + \|x - y\| &\geq \|y\| \\ \|x - y\| &\geq \|y\| - \|x\| \end{aligned}$$

and

$$\begin{aligned} \|y\| + \|y - x\| &\geq \|x\| \\ \|y - x\| &\geq \|x\| - \|y\| . \end{aligned}$$

Useful inequalities

Theorem (Cauchy-Schwarz inequality)

Let V be an inner product space and let $u, v \in V$. Then,

$$|\langle u, v\rangle| \leq \|u\| \, \|v\| \, .$$

Proof.

Setup as before, we can show that $t = \frac{\langle u,v\rangle}{\|v\|^2}$. Now, let $z = u - tv$. By the Pythagorean theorem,

$$\begin{aligned}\|u\|^2 &= \|tv\|^2 + \|z\|^2 \\ &= \frac{\langle u, v\rangle^2}{\|v\|^2} + \|z\|^2\end{aligned}$$

$\|z\|^2 \geq 0$, so $\|u\|^2 \geq \frac{\langle u,v\rangle^2}{\|v\|^2}$ $\|u\| \, \|v\| \geq |\langle u, v\rangle|$. □

Section 3

Projections

Projections

Definition

Let V be an inner product space and $x, y \in V$. The **projection** of y onto x is

$$P_x y = \frac{\langle y, x \rangle}{\|x\|^2} x.$$

More generally, the projection of y onto a finite set $\{x_1, x_2, ..., x_k\}$ is

$$P_{\{x_j\}_{j=1^k}} y = \sum_{j=1}^{k} P_{x_j - \sum_{i=1}^{j-1} P_{x_i} x_j} y.$$

Projections

Definition

More generally still, if $X \subseteq V$ is a linear subspace, then the projection of y onto X is

$$P_X y = P_{\{b_j\}_{j=1}^k} y$$

where $b_j \in X$ and $b_1, ..., b_k$ span X.
Finally, if $Y \subseteq V$ the projection of Y onto X is just the set consisting of the projection of each element of y onto X, i.e.

$$P_X Y = \{P_X y : y \in Y\}.$$

Projections

Lemma

Any projection is an idempotent linear transformation.

Proof.

- Verify that projections have the two properties required for them to be linear transformations.
- Show that projections are idempotent.

$$
\begin{aligned}
P_x(P_x y) = & P_x \left(\frac{\langle x, y \rangle}{\|x\|^2} x \right) \\
= & \frac{\left\langle x, \frac{\langle x, y \rangle}{\|x\|^2} x \right\rangle}{\|x\|^2} x \\
= & \frac{\langle x, y \rangle}{\|x\|^2} \frac{\langle x, x \rangle}{\|x\|^2} x \\
= & P_x y.
\end{aligned}
$$

Section 4

Row, column, and null space

Row space

Definition

Let A be an m by n matrix. The **row space** of A, denoted $\mathrm{Row}(A)$, is the space spanned by the row vectors of A.

- $\mathrm{Row}(A) \subseteq \mathbb{R}^n$

Lemma

Performing Gaussian elimination does not change the row space of a matrix.

Proof.

Let $a_1, ..., a_m$ be the row vectors of A. Each step of Gaussian elimination transforms some a_j into $a_j + g a_k$ with $k \neq j$ or $g \neq -1$. Can show that

$$\mathrm{span}(a_1, ..., a_m) = \mathrm{span}(a_1 + g a_k, ..., a_m).$$

□

Corollary

The dimension of the row space of a matrix is equal to its rank.

Column space

Definition

Let A be an m by n matrix. The **column space** of A, denoted $\mathrm{Col}(A)$, is the space spanned by the column vectors of A.

- $\mathrm{Col} A \subseteq \mathbb{R}^m$

Lemma

Let A be an m by n matrix. Then $Ax = b$ has a solution iff $b \in \mathrm{Col}(A)$.

Definition

A column of a matrix, A, is **basic** if the corresponding column of the row echelon form, A_r, contains a pivot.

Theorem

The basic columns of A form a basis for $\mathrm{Col}(A)$.

Proof.

Let A be $m \times n$ and denote its columns as $v_1, ..., v_n$. Let A_r be the row echelon form of A and denotes its columns as $w_1, ..., w_n$. Let $w_{i_1}, ..., w_{i_k}$ be the basic columns of A_r. Each has more zeros, so $w_{i_1}, ..., w_{i_k}$ are linearly independent. By definition of row echelon form, the final $m - k$ rows of A_r are all zero. Therefore $\dim \mathrm{Col}(A_r) \leq k$, and $w_{i_1}, ..., w_{i_k}$ must be a basis for $\mathrm{Col}(A_r)$. □

Continued.

Now we show that $v_{i_1}, ..., v_{i_k}$ are a basis for $\mathrm{Col}(A)$. Suppose

$$c_1 v_{i_1} + ... + c_k v_{i_k} = 0.$$

Then we could do Gaussian elimination to convert this system to

$$c_1 w_{i_1} + ... + c_k w_{i_k} = 0.$$

$w_{i_1}, ..., w_{i_k}$ are linearly independent so $c_1 = 0, ... c_k = 0$. Add any other v_j, $j \notin \{i_1, ..., i_k\}$, then by the same argument there must exist a non-zero c than solves

$$c_1 v_{i_1} + ... + c_k v_{i_k} + c_j v_j = 0.$$

Thus, $v_{i_1}, ..., v_{i_k}$ is a basis for $\mathrm{Col}(A)$. □

Corollary

The dimensions of the row and column spaces of any matrix are equal.

Corollary

$\mathrm{rank} A = \mathrm{rank} A^T$.

Null space

Definition

Let A be m by n. The set of solutions to the homogeneous equation $Ax = 0$ is the **null space** (or kernel) of A, denoted by $\mathcal{N}(A)$ (or $\mathrm{Null}(A)$).

Definition

Let $V \subseteq \mathbb{R}^n$ be a linear subspace, and let $c \in \mathbb{R}^n$ be a fixed vector. The set

$$\{x \in \mathbb{R}^n : x = v + c \text{ for some } v \in V\}$$

is called the set of **translates** of V by c, and is denoted $c + V$. Any set of translates of a linear subspace is called an **affine** space.

Lemma

Let $Ax = b$ be an m by n system of linear equations. Let x_0 be any particular solution. Then the set of solutions is $x_0 + \mathcal{N}(A)$.

Proof.

Let $w \in x_0 + \mathcal{N}(A)$. Then

$$\begin{aligned} Aw =& A(x_0) + A(\underbrace{w - x_0}_{\in \mathcal{N}(A)}) \\ =& b + 0. \end{aligned}$$

Let w be a solution to $Ax = b$. Then

$$A(w - x_0) = Aw - Ax_0 = 0$$

so $w - x_0 \in \mathcal{N}(A)$ and $w \in x_0 + \mathcal{N}(A)$. □

Theorem

Let A be an m by n matrix. Then $\dim \mathcal{N}(A) = n - \operatorname{rank} A$

Proof.

- Let $u_1, ..., u_k$ be a basis for $\mathcal{N}(A)$. We can add $u_{k+1}, ..., u_n$ to $u_1, ..., u_k$ to form a basis for $\mathbb{R}^n$.
- Show that $Au_{k+1}, ..., Au_n$ are a basis for the column space
  - linearly independent
  - span $\operatorname{Col} A$.

□

Theorem (Rouché-Capelli)

A system of linear equations with n variables has a solution if and only if the rank of its coefficient matrix, A, is equal to the rank of its augmented matrix, $\hat{A}$. Equivalently, a solution exists if and only if $b \in \mathrm{Col}(A)$.
If a solution exists and $\mathrm{rank} A$ is equal to its number of columns, the solution is unique. If a solution exists and $\mathrm{rank} A$ is less than its number of columns, there are infinite solutions. In this case the set of solutions forms is $x_0 + \mathcal{N}(A)$, where x_0 is any particular solution to $Ax = b$. This set of solutions is an affine subspace of dimension $n - \mathrm{rank} A$.

Relationship among row, column, and null spaces

- $\mathrm{Col}(A) = \mathrm{Row}(A^T) \subseteq \mathbb{R}^m$
- $\mathrm{Row}(A) = \mathrm{Col}(A^T) \subseteq \mathbb{R}^n$
- $\mathcal{N}(A) \subseteq \mathbb{R}^n$ and $\mathcal{N}(A^T) \subseteq \mathbb{R}^m$
- Let $x \in \mathcal{N}(A)$, $y \in \mathrm{Row}(A)$, what is $\langle x, y \rangle$?

Relationship among row, column, and null spaces

- If $x \in \mathcal{N}(A)$, $y \in \mathrm{Row}(A) = \mathrm{Col}(A^T)$, then $\langle x, y \rangle = 0$.
- $\mathcal{N}(A)$ and $\mathrm{Row}(A)$ are orthogonal subspaces of $\mathbb{R}^n$
- If $x \in \mathcal{N}(A^T)$, $y \in \mathrm{Row}(A^T) = \mathrm{Col}(A)$, then $\langle x, y \rangle = 0$.
- $\mathcal{N}(A^T)$ and $\mathrm{Col}(A)$ are orthogonal subspaces of $\mathbb{R}^m$
- $\forall x \in \mathbb{R}^n$,

$$\begin{aligned} Ax = & A(P_{\mathrm{Row}A} x + P_{\mathcal{N}(A)} x) \\ = & A(P_{\mathrm{Row}A} x) + A(P_{\mathcal{N}(A)} x) \\ = & A(P_{\mathrm{Row}A}) x \in \mathrm{Col}(A) = \mathrm{Row}(A^T) \end{aligned}$$

- $\forall w \in \mathbb{R}^m$,

$$\begin{aligned} A^T w = & A^T (P_{\mathrm{Col}A} w + P_{\mathcal{N}(A^T)} w) \\ = & A^T (P_{\mathrm{Col}A} w) + A^T (P_{\mathcal{N}(A^T)} w) \\ = & A^T (P_{\mathrm{Col}A}) w \in \mathrm{Row}(A) = \mathrm{Col}(A^T) \end{aligned}$$