Math 484: Nonlinear Programming[1] Mikhail Lavrov

Chapter 1, Lecture 2: Geometry of $\mathbb{R}^n$

January 16, 2019 *University of Illinois at Urbana-Champaign*

1 Motivation

Our next goal is to extend the first and second derivative tests to functions of several variables. There are several questions to ask:

- What is a critical point?
- What are first and second derivatives like in higher dimensions?
- What is the equivalent of the condition $f''(x) > 0$?

But first, we will need to introduce the objects that we'll be dealing with in higher dimensions.

2 Vector arithmetic

Mostly, we think of points in $\mathbb{R}^n$ as column vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ which we can add componentwise and multiply by scalars (real numbers):

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \qquad \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

Often we also think of vectors as arrows with a direction and magnitude. The vector $\mathbf{x}$ corresponds to the arrow going from the origin $\mathbf{0}$ (the all-zero vector) to the point $\mathbf{x}$.

For every column vector $\mathbf{x}$ there is a row vector $\mathbf{x}^\mathsf{T} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$: the transpose of $\mathbf{x}$.

You should think of row and column vectors as being slightly different, much like a volume in cubic centimeters is different from a volume in gallons. You can convert from one to the other, but there's probably a good reason to use one and not the other, so you should be careful to avoid mixing them unnecessarily.

That being the case, sometimes people are lazy and just write $(x_1, x_2, \dots, x_n)$ for both kinds of vectors.

[1] This document comes from the Math 484 course webpage: https://faculty.math.illinois.edu/~mlavrov/courses/484-spring-2019.html

We can multiply an n-dimensional row vector by an n-dimensional column vector and get a real number:

$$\begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n.$$

So if $\mathbf{x}, \mathbf{y}$ are both column vectors—our default representation of points in $\mathbb{R}^n$—we need to take a transpose to multiply them.[2] We write

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\mathsf{T} \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$$

and call this the dot product of $\mathbf{x}$ and $\mathbf{y}$.

3 Matrix arithmetic

Vectors generalize to matrices: an $n \times m$ matrix is just an $n \times m$ grid of numbers. Once again, we have two straightforward operations:

- Two $n \times m$ matrices can be added together, by adding together corresponding entries.
- We can multiply a matrix by a real number, by multiplying each entry by that number.

Then we have the very complicated operation called matrix multiplication. If A is a $k \times n$ matrix and B is an $n \times m$ matrix, then their product AB is a $k \times m$ matrix: its entries are a multiplication table of the *rows* of A by the *columns* of B. That is:

$$AB = \begin{bmatrix} \boxed{\mathbf{A}_1} \\ \boxed{\mathbf{A}_2} \\ \vdots \\ \boxed{\mathbf{A}_k} \end{bmatrix} \begin{bmatrix} \boxed{\mathbf{b}_1} & \boxed{\mathbf{b}_2} & \dots & \boxed{\mathbf{b}_m} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1\mathbf{b}_1 & \mathbf{A}_1\mathbf{b}_2 & \cdots & \mathbf{A}_1\mathbf{b}_m \\ \mathbf{A}_2\mathbf{b}_1 & \mathbf{A}_2\mathbf{b}_2 & \cdots & \mathbf{A}_2\mathbf{b}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_k\mathbf{b}_1 & \mathbf{A}_k\mathbf{b}_2 & \cdots & \mathbf{A}_k\mathbf{b}_m \end{bmatrix}.$$

An example follows a bit further down.

You might often see this written concisely as the formula

$$(AB)_{ij} = \sum_{k=1}^{n} \mathbf{A}_{ik} \mathbf{B}_{kj}$$

(where X_{ij} denotes the $(i,j)^{\text{th}}$ entry of X). This concise formula isn't very useful for understanding how matrices work intuitively. But it's going to be very useful for spotting matrices in the wild: whenever you see a complicated expression with sums that look like this one, you should try to turn it into a simple expression with matrix multiplication.

(In fact, this is the primary way that matrices appear in this class: complicated expressions with sums will appear, and we will turn them into simple expressions involving matrices.)

[2] This might seem like it goes against my earlier warning to avoid mixing row and column vectors, but I promise I am being careful here.

3.1 Example of matrix multiplication

Here is an example of matrix multiplication: we want to multiply a 2×3 matrix by a 3×2 matrix, getting a 2×2 result.

A convenient way to compute the product

$$AB = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

is to write A to the left and B above the final answer. Then to find each entry of the product, you multiply the row vector directly to its left by the column vector directly above it.

Step 1:

$$\begin{matrix} & \begin{bmatrix} \mathbf{1} & 2 \\ \mathbf{3} & 4 \\ \mathbf{5} & 6 \end{bmatrix} \\ \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{1} \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} \mathbf{6} & ? \\ ? & ? \end{bmatrix} \end{matrix}$$

Step 2:

$$\begin{matrix} & \begin{bmatrix} 1 & \mathbf{2} \\ 3 & \mathbf{4} \\ 5 & \mathbf{6} \end{bmatrix} \\ \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{1} \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 6 & \mathbf{8} \\ ? & ? \end{bmatrix} \end{matrix}$$

Step 3:

$$\begin{matrix} & \begin{bmatrix} \mathbf{1} & 2 \\ \mathbf{3} & 4 \\ \mathbf{5} & 6 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 1 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \end{bmatrix} & \begin{bmatrix} 6 & 8 \\ \mathbf{3} & ? \end{bmatrix} \end{matrix}$$

Step 4:

$$\begin{matrix} & \begin{bmatrix} 1 & \mathbf{2} \\ 3 & \mathbf{4} \\ 5 & \mathbf{6} \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 1 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \end{bmatrix} & \begin{bmatrix} 6 & 8 \\ 3 & \mathbf{4} \end{bmatrix} \end{matrix}$$

4 More on dot products

We define the norm or length of a vector $\mathbf{x} \in \mathbb{R}^n$ as

$$\|\mathbf{x}\| = (\mathbf{x} \cdot \mathbf{x})^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

This is what we use to measure distance in $\mathbb{R}^n$: the distance between $\mathbf{x}$ and $\mathbf{y}$ is

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

There are several fundamental properties of inner products and norms, which are often easier to use directly rather than expanding out the definition and working with the entries of the vectors.

- Linearity of inner product: $(\alpha\mathbf{x} + \beta\mathbf{y}) \cdot \mathbf{z} = \alpha(\mathbf{x} \cdot \mathbf{z}) + \beta(\mathbf{y} \cdot \mathbf{z})$.

  In particular, $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$, and $(\alpha\mathbf{x}) \cdot \mathbf{y} = \alpha(\mathbf{x} \cdot \mathbf{y})$.

- Symmetry of inner product: $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.

  This lets us apply linearity to the second vector in the dot product as well as the first.

- Positivity of inner product: $\mathbf{x} \cdot \mathbf{x} \geq 0$, and $\mathbf{x} \cdot \mathbf{x} = 0$ only if $\mathbf{x}$ is the zero vector $\mathbf{0}$.

  This lets us define the norm $\|\mathbf{x}\| = (\mathbf{x} \cdot \mathbf{x})^{1/2}$ to begin with (without encountering complex numbers).

- Linearity of norm: $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$.
- Triangle inequality: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.
- Cauchy–Schwarz inequality: $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$.

This last inequality is used a lot in general, but won't come up much in this class. It's related to how we compute angles between vectors: the angle θ between vectors $\mathbf{x}$ and $\mathbf{y}$ (thinking of $\mathbf{x}$ and $\mathbf{y}$ as arrows, not as points) satisfies

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\| \cos \theta.$$

This might be a theorem, or a definition of angles, depending on how you go about doing things.

In particular—this is the important consequence—when $\mathbf{x} \cdot \mathbf{y} = 0$, we conclude that $\cos \theta = 0$, so $\theta = 90^\circ$: vectors whose dot product is 0 are perpendicular. (Or, in linear algebra terminology, orthogonal.)