ECON 4140 Mathematics 3 spring-18

Lecture 1: preliminaries (updated)

Will use vector/matrix notation:

$\vec{x}$ or $\bar{x}$ or $\mathbf{x}$ or $\underset{\frown}{x}$ or $x \in \mathbb{R}^n$

$\vec{A}$ or $\bar{A}$ or $\mathbb{A}$ ... your preference? | You chose $\vec{x}$, $\vec{A}$

Vectors are by default columns [But: "(!)" below]

Row: $\vec{x}^T$ (book uses $\mathbf{x}'$)

$f(\vec{x})$: function of n variables
$(n \geq 1)$ $f(x_1, \ldots, x_n)$ as if x were row (!)

Notation: $\nabla f(\vec{x}) = \left(\frac{\partial f}{\partial x_1}(\vec{x}), \ldots \frac{\partial f}{\partial x_n}(\vec{x})\right)$ row (!)

(the "gradient"): vector of 1^{st} partial deriv's.

Also: Hessian matrix $\vec{H} = \vec{H}(x)$:
$\vec{H} = (h_{ij})$, $h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x})$ of
second derivatives.

Note: "Hessian" can mean $\vec{H}$ or $|\vec{H}|$

Can write a F.O.C as
$$\nabla f(\vec{x}) = \vec{0}$$
(!) – really $= \vec{0}^T$

... and: $\nabla f(\vec{x}) = \sum_{j=1}^{m} \lambda_j \nabla g_j(\vec{x})$

Transformations: term the book uses for "functions that return other things than numbers" (e.g.: vectors, matrices, functions, sets).

* Most Math 3 - relevant: $\vec{F}(\vec{x}) = \begin{pmatrix} F_1(\vec{x}) \\ \vdots \\ F_m(\vec{x}) \end{pmatrix}$

For such "vector valued functions":
the matrix $\begin{pmatrix} \nabla F_1(\vec{x}) \\ \vdots \\ \nabla F_m(\vec{x}) \end{pmatrix}$ of partial 1^{st} derivatives

is called the Jacobian.
(Do not confuse with Hessian!)

* Other transformations: examples:

	input	output
$\frac{\partial}{\partial x}$	function	function
∇	function	row vector of functions
Hessian	function	matrix of functions

and the following: given preferences, let
$B(\vec{x})$ = the set of all $\vec{z}$ that are (weakly) preferred to $\vec{x}$

* Using linear algebra notation in analysis has some issues....

I will often use

$\vec{x}^T$ for transposition, to distinguish from derivative

$\det \vec{A}$ or $\det(\vec{A})$ for determinant, to distinguish from absolute value of a number.

$\vec{v}^{(j)}$ for vector number j (not component)

* Q: How do differentiation rules look with linear algebra?

A: Later! Only a couple of basic examples for now:

$f(\vec{x}) = \vec{p}^T\vec{x}$ (equals $\vec{x}^T\vec{p}$) has $\nabla f(x) = \vec{p}^T$ and Hessian $= \vec{0}_{n\times n}$

$g(\vec{x}) = \vec{x}^T\vec{A}\vec{x}$ (equals, in fact, $\frac{1}{2}\vec{x}^T(\vec{A}+\vec{A}^T)\vec{x}$)

$\nabla g(\vec{x}) = \vec{x}^T(\vec{A}+\vec{A}^T)$, Hessian $= \vec{A}+\vec{A}^T$

Note: if $\vec{A}=\vec{I}$ then $g(\vec{x}) = \|\vec{x}\|^2$, $\|\vec{x}\|$ = Euclidean norm (= "length"),

(Differentiating $\|\vec{x}\|$ is much worse!

Fact: $\nabla \|\vec{x}\| = \frac{1}{\|\vec{x}\|}\vec{x}$, Hessian $= \frac{1}{\|\vec{x}\|}\vec{I} - \frac{1}{\|\vec{x}\|^3}\vec{x}\vec{x}^T$)

Lecture 1 cont'd; linear algebra

<u>Def.</u>: A <u>minor</u> of a matrix $\vec{A}$ is a determinant $|\vec{M}|$ – or the value of such one – formed by deleting 0 or more columns and 0 or more rows from $\vec{A}$, so the resulting matrix $\vec{M}$ is square

<u>Ex.</u> $\vec{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ has three 2×2 minors

$\left(\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}, \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \text{ and } \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \right)$

and six 1×1 minors (the elements).

<u>Note</u>: In other literature, you can find "minor" for the <u>matrix</u> $\vec{M}$, and/or the convention that "<u>something</u> must be deleted".

In this course, $|\vec{A}|$ is a minor of $\vec{A}$ as long as $\vec{A}$ is square.

Linear combinations

Let $\vec{v}^{(1)}, \dots \vec{v}^{(n)}$ be vectors in the same vector space (say, $\mathbb{R}^m$).

A linear combination of these vectors is a sum $c_1 \vec{v}^{(1)} + \dots + c_n \vec{v}^{(n)}$

for some numbers $c_1, \dots, c_n$

Note: the concept works beyond $\mathbb{R}^m$.

Ex: $c_1 \bar{A}_1 + \dots + c_n \bar{A}_n$ for matrices

$c_1 f_1 + \dots + c_n f_n$ functions

$c_1 X_1 + \dots + c_n X_n$ random var's

and more.

Other terminology:

- the linear span of $\{\vec{v}^{(1)}, \dots, \vec{v}^{(n)}\}$: the set of possible $\vec{w}$ that can be written as $\vec{w} = c_1 \vec{v}^{(1)} + \dots + c_n \vec{v}^{(n)}$
- Convex combination: if we restrict the c_i to satisfy $c_i \geq 0$, $c_1 + \dots + c_n = 1$.

[Weighted average, possibly zero weights]

Linear (in) dependence

Def. a set S of vectors $\{\vec{v}^{(1)}, \dots, \vec{v}^{(n)}\}$
(in the same space, say, $\mathbb{R}^m$)

is called linearly dependent if

there exist $c_1, \dots, c_n$ not all zero,

such that

$$c_1 \vec{v}^{(1)} + \dots + c_n \vec{v}^{(n)} = \vec{0} \qquad (*)$$

and linearly independent if $(*)$ only holds when $c_1 = c_2 = \dots = c_n = 0$.

Ex.: Linearly independent: $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$

(For $(*)$ to hold, c_1 must be $= -c_2$, yields $c_2 \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ so $c_2 = 0$, and $c_1 = -0 = 0$.)

Linearly dependent: $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \right\}$

(For example, put $c_1 = -3$, $c_2 = 0$, $c_3 = -1$, $c_4 = 1$.)

Note: property of the set!

(The king & I are not brothers just because he is a brother and I am one.)

Properties:

- If $\vec{0} \in S$ then S is linearly dependent
  (put all other $c_i = 0$)

(: $\emptyset$ is lin. indep)

- A singleton $\{\vec{v}\}$ is linearly dependent
  only if $\vec{v} = \vec{0}$.

- A pair $\{\vec{u}, \vec{v}\}$ is linearly dependent
  only if they are colinear (= "proportional"
  where $\vec{0}$ is "proportional to anything")

- A triplet $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly dependent
  only if they are coplanar.

- Remove a vector from a linearly
  independent set $\rightsquigarrow$ a lin. independent set.

- Augment a linearly dependent set with
  more vector(s) $\rightsquigarrow$ a linearly dependent set

- If $S \subseteq \mathbb{R}^m$ and has $> m$ elements,
  then S must be linearly dependent!

Why the latter? With n lin. indep. vectors on $\mathbb{R}^n$, we have

$$c_1 \vec{v}^{(1)} + \ldots + c_n \vec{v}^{(n)} = \underset{\uparrow_{n\times n}}{\vec{V}}\, \vec{c} \quad \text{with}$$

$\vec{V} = (\vec{v}^{(1)} \vdots \cdots \vdots \vec{v}^{(n)})$. Now:

lin indep

$\Updownarrow$

the eq. $\vec{V}\vec{c} = \vec{0}$ has unique solution $\vec{c} = \vec{0}$

$\Updownarrow$

$\vec{V}^{-1}$ exists (since $\vec{V}$ square!)

$\Updownarrow$

the eq $\vec{V}\vec{c} = \vec{B}$ has unique solution $\vec{c} = \vec{V}^{-1}\vec{B}$.

Introduce one more vector $\vec{b}$, put its coefficient, c_{n+1} equal to -1:

$$\vec{0} = c_1 \vec{v}^{(1)} + \ldots + c_n \vec{v}^{(n)} + (-1)\vec{b}$$

$$\Leftrightarrow \quad \vec{V}\vec{c} = \vec{b} \quad \Leftrightarrow \quad \vec{c} = \vec{V}^{-1}\vec{b}.$$

(With more vectors $\vec{B}$ similar!)

So <u>even when</u> we have linear independence with equally many vectors as coordinates, then increasing the number of vectors <u>must</u> introduce linear dependence!