MA1C ANALYTIC RECITATION 5/10/12

ALDEN WALKER

1. AES

I am not sure whether you learned about vector spaces over fields other than $\mathbb{R}$ in math 1b. I thought you might be interested to know about the Advanced Encryption Standard (AES). Notice that the set $\{0, 1\}$ under regular multiplication and addition mod 2 is a field (often called GF(2)). Therefore, we can consider vector spaces over this field, and we can do linear algebra. Everything you learned in 1b is still true for linear algebra over GF(2), except stuff with eigenvalues can be weird.

Among all computers in the world doing linear algebra right now, by far the majority will be doing linear algebra over GF(2) or another finite field like GF(2^8). One of the reasons for this is cryptographic computations, and in particular AES. If you know about RSA or Diffie-Hellman (public key cryptosystems) you may wonder why there is a need for a regular, symmetric key cryptosystem. The reason is that public key systems tend to be extremely costly computationally. Therefore, they are used for an initial exchange of secret information, from which a key is generated and used in a symmetric system, like AES. The way AES works is the following: the key is 128 bits, and data is encrypted in 128-bit blocks; think of it as a four by four array of bytes. A round of AES consists of doing an S-box, a shiftrows, a mixcolumns, and an addkey which are:

S-box: This is a nonlinear function (usually done by table lookup) on each of the 16 bytes individually.

Shiftrows: In this step, the rows of the 4 by 4 array are shifted to the left by their row index, i.e. the first (zeroth) row is left unchanged, the second row is shifted left one, the third twice, and the fourth 3 times.

Mixcolumns: This is where the linear algebra comes in; each column of the array is thought of as a vector in GF(2^8)^4 and is multiplied by a matrix, the result of which replaces that column.

Addkey: In the final step of the round, the key is added (remember this is mod 2, or xor) to the array bitwise.

To encrypt something in AES, you put your data into the array and run that through 10 rounds as described above. Decryption is basically the same thing, just with inverses.

So that's what happens to your credit card information (hopefully!) when you buy something online.

2. Line Integrals

The motivation for line integrals is that you imagine yourself travelling along inside of a vector field, say a force field (like for gravity). As you go, the field does work on you, and you want to know the total work done over the course of the trip. In general, you want to know "how much" a vector field points along a path. These are actually one-varible integrals, but it doesn't make as much sense to study them in 1a. The definition is as follows: suppose that $\mathbf{c}$ is a piecewise smooth path in $\mathbb{R}^n$, meaning that $\mathbf{c}$ is a continuous map from $[a, b] \subset \mathbb{R}$ to $\mathbb{R}^n$, and there is a partition $[t_0, t_1] \cup [t_1, t_2] \cup \cdots \cup [t_{m-1}, t_m]$ of the interval (ignore those overlapping endpoints) such that $\mathbf{c}$ restricted to any of the intervals in the partition is smooth (infinitely differentiable) (actually just differentiable is all you need). Suppose also that $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is a vector field defined and bounded on the image of $\mathbf{c}$. Then the line integral of $\mathbf{f}$ along $\mathbf{c}$ is written $\int \mathbf{f} \cdot d\mathbf{c}$, and is actually computed as:

$$\int \mathbf{f} \cdot d\mathbf{c} = \int_a^b \mathbf{f}(\mathbf{c}(t)) \cdot \mathbf{c}'(t)\, dt$$

2.1. **Other Notation.** I find the following notation confusing, but I admit it is quite nice for some things. Write $\mathbf{f} = (f_1, \cdots, f_n)$ and $\mathbf{c} = (c_1, \cdots, c_n)$. Then note that the integral is $\int_a^b \sum_i f_i(c(t))c_i'(t)\,dt$, which is written $\int f_1 dc_1 + \cdots + f_n dc_n$. Therefore if you encounter (as you will) a question like: compute the line integral $\int \frac{xdy+ydx}{x+y}$ over the path C, what that means is: define $\mathbf{f}(x,y) = \left(\frac{y}{x+y}, \frac{x}{x+y}\right)$ (note the order!) and take the line integral of that over the path C (you have to parameterize the path).

2.2. **Does Parametrization Matter?** Suppose that I traverse the same path, but I do it at a different speed—will the line integral be different? No. Here is why: what that question is really asking is: suppose that you have some orientation-preserving diffeomorphism h from another interval $[c,d]$ to $[a,b]$: is $\int \mathbf{f}\cdot d\mathbf{c} = \int \mathbf{f}\cdot d(\mathbf{c}\circ h)$. An example would be the map $h : [0,1/2] \to [0,1]$ defined $h(t) = 2t$. Then $\mathbf{c}\circ h : [0,1/2] \to \mathbb{R}^n$ is a path which has the same image as $\mathbf{c}$, but it does it twice as fast.

Let's see that those integrals are the same:

$$\begin{aligned}
\int \mathbf{f}\cdot\, d(\mathbf{c}\circ h) &= \int_c^d \mathbf{f}(\mathbf{c}(h(t)))\cdot(\mathbf{c}\circ h)'(t)\,dt \\
&= \int_c^d \mathbf{f}(\mathbf{c}(h(t)))\cdot\mathbf{c}'(h(t))h'(t)\,dt \\
&= \int_{h(c)}^{h(d)} \mathbf{f}(\mathbf{c}(u))\cdot\mathbf{c}'(u)\,du \\
&= \int \mathbf{f}\cdot\, d\mathbf{c}
\end{aligned}$$

Intuitively, the reason changing parametrizations doesn't change the line integral is that changing the parameterization changes the derivative of your path, which cancels out any change you get in the value of $\mathbf{f}(\mathbf{c}(t))$.

2.3. **Example.** Compute $\int_C -ydx + xdy$ counterclockwise around the circle C of radius 4 in the plane.

First, we need to parameterize the circle: we can do this as $\mathbf{c}(t) = (4\cos(t), 4\sin(t))$, where $t \in [0, 2\pi]$, so $\mathbf{c}'(t) = (-4\sin(t), 4\cos(t))$. The function $\mathbf{f}$ is given as $\mathbf{f}(x,y) = (-y,x)$. Therefore, we compute:

$$\begin{aligned}
\int_C -ydx + xdy &= \int_0^{2\pi} \mathbf{f}(\mathbf{c}(t))\cdot Bc'(t)\,dt \\
&= \int_0^{2\pi} (-4\sin(t), 4\cos(t))\cdot(-4\sin(t), 4\cos(t))\,dt \\
&= \int_0^{2\pi} 16(\sin^2(t)+\cos^2(t))\,dt \\
&= 32\pi
\end{aligned}$$

By the way, you cannot find a function $g : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla g = (-y, x)$. Why is that? Stay tuned for the next section!

3. Line Integrals With Respect to Arc Length

Another common thing you might want to do is to simply integrate a scalar function along some path in n dimensions. That's not really what a line integral computes. However, it is exactly what a line integral with respect to arc length is for. Suppose you have a scalar field $f : \mathbb{R}^n \to \mathbb{R}$ and a path $\mathbf{c}(t) : [0,1] \to \mathbb{R}^n$. Then the line integral with respect to arc length of f along $\mathbf{c}$ is $\int f\,ds = \int_0^1 f(\mathbf{c}(t))\|\mathbf{c}'(t)\|\,dt$. Essentially, this is just like doing $\int f dx$, where the dx represents an a little piece of distance. Here, the path might speed up and slow down or not go at speed one, so the right thing to stick it to get a little piece of distance is $\|\mathbf{c}'(t)\|\,dt$. These integrals tend to be harder to compute because of the annoying square root; therefore, if you can pick a path with constant speed, you will thank yourself.

Note that a line integral with respect to arc length is more general than a line integral: if you take the scalar function $f(c(t)) \cdot c'(t)/\|c'(t)\|$, then notice you recover the definition of a line integral.

3.1. **Example.** Let's integrate the scalar function $f(x, y) = x^2$ around the circle $c(t) = (\cos t, \sin t)$. Here we have $\|c'(t)\| = 1$, so our integral is:

$$\int f\, ds = \int_0^{2\pi} 1(\cos^2 t)dt = \left(\frac{t}{2} + \frac{\sin(2t)}{4}\right|_0^{2\pi} = \pi$$

You can imagine this as the area above the xy-plane of a sheet suspended from the graph of the parametric function $f(c(t)) = (\cos t, \sin t, \cos^2 t)$.

4. Gradient Fields

You already know a great way to make vector fields in $\mathbb{R}^n$—just get a scalar function and take its gradient. Well, gradient fields have some very nice properties with respect to line integrals, first, there is independence of path. Specifically, Theorem 10.3, which says that if φ is a differentiable scalar field with a continuous gradient on an open connected set, then for any path $\mathbf{c}$ with endpoints $\mathbf{c}(0) = a$ and $\mathbf{c}(1) = b$ we have $\int_0^1 \nabla\varphi \cdot d\mathbf{c} = \varphi(b) - \varphi(b)$. This is the case because the function $\varphi(\mathbf{c}(t))$ is a map $\mathbb{R} \to \mathbb{R}$ with (one-variable) derivative $\nabla\varphi(\mathbf{c}(t)) \cdot \mathbf{c}'(t)$, so we apply the fundamental theorem of calculus.

Path independence is very nice, since to calculate a line integral you can just choose whatever path you want! It also implies that the integral around any closed loop is zero (why?). It turns out that the following are equivalent (Theorem 10.5):

(1) The vector field F is a gradient field.
(2) The line integral of F around any closed loop is zero.
(3) The line integral of F depends only on the endpoints of the path.

4.1. **Example.** Take the scalar function $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(x, y) = (x+y)^2$. Then $\nabla f = 2(x+y, x+y)$. What is the line integral of ∇f from $(-1, 1)$ to $(1, 1)$?

We can just use the facts above to see that $\int_C \nabla f \cdot d\mathbf{c} = f(1, 1) - f(-1, 1) = 4$, but let's do it out to make sure. Define $\mathbf{c}(t) = (t, 1)$ for $t \in [-1, 1]$. Then

$$\begin{aligned}
\int f \cdot d\mathbf{c} &= \int_{-1}^{1} \nabla f(\mathbf{c}(t)) \cdot \mathbf{c}'(t)dt \\
&= \int_{-1}^{1} 2(t+1, t+1) \cdot (1, 0)dt \\
&= 2\int_{-1}^{1} t + 1dt \\
&= 2\left(\frac{t^2}{2} + t\right|_{-1}^{1} \\
&= 2\left(\frac{1}{2} + 1 - \frac{1}{2} + 1\right) \\
&= 4
\end{aligned}$$

But why not let's do it again! Now set $\mathbf{c}(t) = (t, t^2)$, so $\mathbf{c}'(t) = (1, 2t)$. Then

$$\begin{aligned}
\int f \cdot d\mathbf{c} &= \int_{-1}^{1} \nabla f(\mathbf{c}(t)) \cdot \mathbf{c}'(t) dt \\
&= \int_{-1}^{1} 2(t + t^2, t + t^2) \cdot (1, 2t) dt \\
&= 2\int_{-1}^{1} t + t^2 + 2t^2 + 2t^3 dt \\
&= 2\left(\frac{t^2}{2} + t^3 + \frac{t^4}{2}\right|_{-1}^{1} \\
&= 2\left(\frac{1}{2} + 1 + \frac{1}{2} - \frac{1}{2} + 1 - \frac{1}{2}\right) \\
&= 4
\end{aligned}$$

So yup that checks out.

It matters that the set is connected! It also matters what the set *is*. For example, you will show on your homework that the vector field $F(x, y) = (-y/(x^2+y^2), x/(x^2+y^2))$ is the gradient of a function on $\mathbb{R}^2$ with the negative x axis removed. However, this function is *not* the gradient of a function on all of $\mathbb{R}^2$. Why? Observe that

$$\int_C F \cdot ds = 2\pi \neq 0$$

Where C is the unit circle counterclockwise.