Miscellaneous Notes

Matrix Exp. Let $X(t)$ be a fundamental matrix of a homogeneous linear d.e. with coefficient matrix $A(t)$,*not necessarily constant*. Note that $X(t)$ is well-defined up to an invertible factor on the right; hence $X(t)X(0)^{-1}$ is independent of the choice of $X(t)$; we shall denote it by $X_A(t)$.

$X_A(t)$ is characterized by $\dot{X}_A = AX_A$ and $X_A(0) = I$. If A is constant , $X_A(t) = expAt$.

Let B be another matrix.The identity

$$\frac{d}{dt}(X_A X_B) = (A + B)X_A X_B + (X_A B - BX_A)X_B$$

shows that $X_A X_B = X_{A+B}$ iff B commutes with X_A. If A and B are constant, this means that $AB = BA$ (look at terms of degree 1 in the power series).

NOTE: If A is not constant, it does not define a vector-field, and $X_A(t)$ is not a flow in the usual sense.

Binomial Trials and Difference Equations. By a *bell-curve of width σ centered at μ*, we mean a positive solution of the differential equation

$$\frac{dy}{dx} = -\frac{x - \mu}{\sigma^2} y.$$

Fix a natural number n and positive reals p, q with $p + q = 1$. On the interval $[-1, n + 1]$, consider the piece- wise linear function $b(x)$ which is 0 at the end- points and has

$$b(k) = \binom{n}{k} p^k q^{n-k}$$

at integers k in between. On each subinterval $(k, k + 1)$, its slope $b(k + 1) - b(k)$ turns out to be very easily computed in terms of the intermediate point $x_k = q(k + 1) + pk$, namely

$$b(k + 1) - b(k) = -\frac{x_k - np}{(n + 1)pq} b(x_k).$$

Thus, on the discrete set of points $(x_k, b(x_k))$, the graph of $b(x)$ satisfies the differential equation for bell-curves of width $\sigma_n = \sqrt{(n + 1)pq}$ centered at $\mu_n = np$.

For a picture less dependent on n, we transfer our graph to a (u, v)-plane, with $u = (x - \mu_n)/\sigma_n$ and $v = \sigma_n y$. The points $(x_k, b(x_k))$ go into (u_k, v_k) at which the image graph now has slope $-u_k v_k$, i.e. looks like a bell curve of width 1 centered at 0, its support reaching from $-(np + 1)/\sigma_n$ to $(nq + 1)/\sigma_n$. For large n it therefore seems (!) that the graph can be reconstructed from its apex —using Stirling's formula to get $b([n/2])$— by solving the linear d.e. $v' = -uv$ via finite differences.

Remark: Of course we are not seriously advocating such a reconstruction, for which there would be the exact (but non-linear) formula $b(k+1) - b(k) = c(k)b(k)$, where $c(k) = ((n+1)p - (k+1))/(k+1)q$. Rather, we are trying to show an a priori proximity between binomial and normal distribution. The unusual +1 in our formula for σ_n probably comes from working with the piece-wise linear function b instead of the usual histogram.

Musical Mnemonic. Succesive intervals of the diatonic scale are:

$$\alpha, \beta, \epsilon, \alpha, \beta, \alpha, \epsilon$$

, where $\alpha = 1 + 1/8, \beta = 1 + 1/9, \epsilon = 1 + 1/15$. The rest of the drama follows from this. Note, for instance: however you define F-sharp, the scale of G gets off to a bad start. Even worse, the interval from D to A is $\alpha\beta^2\epsilon$ instead of the perfect fifth $\alpha^2\beta\epsilon$. The twelfth root of 2 was applied to this puzzle in 1686 by one Werckmeister. A previous expert had come close by using 18/17.

The Gradient of a Quadratic Form. Let A be a linear transformation on a Eucidean space and put $Q(X) = AX \bullet X$. The identity $Q(X+V) - Q(X) = (A + A^t)X \bullet V + Q(V)$ clearly shows that

$$\nabla Q(X) = (A + A^t)X.$$

For those who are not comfortable with Frechet derivatives, we could use the same identity to compute the directional derivative $D_U Q(X) = \lim_{h\to 0}[Q(X+hU) - Q(X)]/h = (A+A^t)X \bullet U$. Either way it is immediate that, for *symmetric* A, the problem of maximizing $Q(X)$ on the unit sphere (using Lagrange multipliers) is identical to the eigenvalue problem.

Combinations with Repeats. According to Melzak the following pretty argument is due to Euler. If $S_k = \{1, \dots, k\}$, let $F(n,k)$ and $G(n,k)$ stand for the increasing and non-decreasing functions $S_k \to S_n$, respectively. The cardinality of the former is good old $C(n,k)$; what is the cardinality of $G(n,k)$, the set of "combinations with repeats"?

Given $g \in G(n,k)$, construct a function $f \in F(n+k-1,k)$ by making $f(x) = g(x) + x - 1$. Conversely, given such an f, put $g(x) = f(x) - x + 1$; then $g(x+1) = f(x+1) - x > f(x) - x$, i.e. $g(x+1) \geq g(x)$. Hence $G(n,k)$ is in one-to-one correspondence with $F(n+k-1,k)$.

The Cross Product. For $V, W \in \mathbf{R}^3$, define $V \times W$ by

$$(V \times W) \bullet X = \det(V, W, X) \qquad \text{for all} \quad X \in \mathbf{R}^3.$$

Then it is obvious that

$$A^t(AV \times AW) = \det A \cdot (V \times W),$$

for any 3×3-matrix A. Hence the product is invariant under rotations.

Moreover it is clear that $V \times W$ is orthogonal to V, W. To interpret its length, let U be a unit vector parallel to it. Then $|V \times W| = |(V \times W) \bullet U| = |\det(V, W, U)|$, which equals the area of the parallelogram given by V, W.

The Simplex Method. A linear optimization problem involves a (consistent) system $AX = B$ of linear equations and a linear pay-off function C on the solution set thereof. It is in *standard form* if A is fully reduced, B is non-negative, and C is expressed in terms of the "free" variables. The simplex method treats such a problem by moving from one standard form to an adjacent one while making sure that C is increased. The following three steps are iterated.

1. Look at the formula for C to decide which one of the free variables (if any) should be "entered".
2. Which row, in the corresponding column, can be used as a pivot without introducing negatives on the right? Check constant-to-coefficient ratios.
3. Sweep out the column chosen in (1) using the row chosen in (2). Don't forget to include (and thus modify) the formula for C in this process. – Go back to (1).

Sion's Lemma. Here is Maurice's version of the singular value decomposition.

Let $A : U \to V$ be linear, and r be the maximum value of $|Au|$ for u on the unit sphere. Put $E = \{u \in U \,|\, Au \bullet Au = r^2 u \bullet u\}$. Then

$$u \in E, \, x \in U \quad \Rightarrow \quad Au \bullet Ax = r^2 u \bullet x.$$

In particular, E is a linear space, and $A(E^\perp) \subseteq A(E)^\perp$.

Proof: Applying Cauchy-Schwarz to the non-negative bilinear form $C(x,y) = r^2 x \bullet y - Ax \bullet Ay$, we get $|C(u,x)|^2 \leq 0$ because $C(u,u) = 0$.

The Möbius Function. The set of functions $f : \mathbf{N} \to D$, where D is any domain, is a ring under pointwise addition and convolution

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d).$$

Its identity is the characteristic function of the set $\{1\}$.

By the fundamental theorem of arithmetic, this ring is isomorphic to the ring of formal power series in countably many indeterminates X_ν (corresponding to prime p_ν). An element f is invertible iff $f(1)$ is a unit in D. Such an f is called *multiplicative* if $f(1) = 1$ and $f = \prod_\nu f_\nu$, where each f_ν is a power series in the single indeterminate X_ν.

The function e which is identically equal to 1, is a case in point. It is the product of geometric series in each of the X_ν; hence its inverse μ, called the *Möbius function*, is the product $\prod_\nu (1 - X_\nu)$. Thus, as a function, $\mu(n) = (-1)^k$ if n has k simple prime factors; if any prime occurs in n with multiplicity > 1, the value of $\mu(n)$ is zero.

Since the equations $g = f * e$ and $f = \mu * g$ are equivalent, we have explicitly

$$g(n) = \sum_{d|n} f(d) \qquad \Longleftrightarrow \qquad f(n) = \sum_{d|n} \mu(d)g(n/d).$$

This is called Möbius inversion.

Dissection Puzzles.

1. Let the vertices of a *regular dodecagon* be numbered as on a clock. Making four cuts from 6 and 7 to 11 and 2, we get six pieces which fit together to form a square. The sides of the latter are the edges joining 7 with 11 and 6 with 2. Compute all relevant angles to check that the fit is perfect (note that 278 is a right angle by Thales).

2. An *equilateral triangle* can be separated into 5 pieces which may be used to form either two or three smaller equilateral triangles. One of the cuts, parallel to the "base", produces a trapezoid; two more, perpendicular to the base, then make a rectangle in which the last cut is a diagonal. The first cut divides two of the sides in the ratio 3:2, with the trapezoid getting the smaller height.

3. What is the shape of an *isosceles triangle* Δ which can be divided into two smaller isosceles triangles by a single cut? Let the legs and base of Δ have lengths 1 and b, and let α and β denote the angles at base and roof, respectively.

If $\alpha > \beta$, the cut must divide one of the base angles, and α remains a base angle for one of the smaller triangles, say Δ', which is therefore similar to Δ. This similarity implies $b^2 = 1 - b$, hence the *golden ratio* for legs to base.

β is the base angle of the other isosceles piece Δ'', and also the roof angle of Δ'. Hence $\alpha = \beta + \beta$. Moreover, the roof angle of Δ'', being the exterior to the base angle of Δ', is $\alpha + \beta = 3\beta$. Summing angles in Δ'', we now get 5β, and thus α is the centre angle of a *regular pentagon*. Therefore Euclid constructs the latter by first making an isosceles triangle based on the golden ratio (which he gets from the diagonal of a rectangular half-square).

For the case $\beta > \alpha$, a similar analysis shows that Δ must have the shape of the Δ'' above.