# Mathematical notes

Neo Quynh Nguyen

December 12, 2020

in due time, this will be completed with more rigorous writting style

### Stochastic fluctuation and viscosity

It is no coincidence that the dimension of kinematic viscosity is $\frac{\text{(length)}^2}{\text{time}}$, while in stochastic calculus and Brownian motion, the mean squared of fluctuation is proportional to time. The diffusion coefficient has the same dimension as kinematic viscosity.

...

### Maximize entropy and likelihood

In logistic regression of a binary classification problem, the logistic function (sigmoid function) $\frac{1}{1+e^{-\theta^T \mathbf{x}}}$ comes from a modeling choice of setting $log\frac{p}{1-p} = \theta^T \mathbf{x}$. This is precisely the Boltzmann probability of being in the lower energy state in a two states system with energy difference $\theta^T \mathbf{x}$. And the distribution comes from the principle of maximum entropy. So what is the connection ?

...

### Maximum likelihood and least square regression

Geometrically, least square method fits a line through the data set such that the sum of the square of the difference between the data and the model is minimize.

The expression of least square loss can be derived by assuming your data is Gaussian distributed, with the mean being your fit model. The log of likelihood will give the least square loss, minus some term.[1]

[1] Bishop, Pattern Recoginition and Machine learning. page 29

is it a surprise that Gaussian distribution and geometrical consideration give the same answer ? not really, many independent data points has Gaussian distribution (Central limit theorem), and the mean is the the likelihood maximum.

...

**Euler La Lagrangian equation, Variance and Jensen's inequality for convex function**

supposed in zero gravity, and in 1 space dimension x and 1 time dimension t. the trajectory of a point particle from (x1,t1) to (x2,t2) in (x,t) space is x(t) curves that minimize the action:

$$S = \int_{t_1}^{t2} (\frac{1}{2}\dot{x})^2 dt$$

S is the mean squared of the velocity, which is always greater or equal the square of mean velocity (Jensen's inequality for convex function say that the expectation

$$E(\phi(x)) \leq \phi(E(x)$$

for convex function $\phi(x)$). so S is minimized when

$$\dot{x} = < x > = \frac{x2 - x1}{t2 - t1}$$

and the trajectory is then

$$x(t) = x1 + t < x >$$

that is, if there is no potential, x is linear in t, and there is no acceleration.

the integrand is called the Lagrangian L, in this case L$=(\frac{1}{2}\dot{x})^2$ =T (kintetic energy). if there is a potential V = gx, then L(x,$\dot{x}, t$)= T-V. the equation of motion is the x*(t) curve that minimizes the action S. T is punishment for large derivative (going to fast), while V is the reward for going to lower potential. i.e. the particle is constrained to go from x1,t1 to x2,t2. at x2 it has lower potential than that of x1, so particle goes there. average speed is fixed as $< x > = \frac{x2-x1}{t2-t1}$ , what path to take ? that path that the straightest in warped space ?.

**macroscopic and microscopic momentum flux and pressure**

physics curriculum nowaday doesn't include much continuum mechanics, thus the idea of momentum flux is unfamiliar to an average physicists. what is the intuition behind it ?

in a fluid, the momentum flux is

$$\rho u^2 + p$$

where $\rho$ is density, u is macroscopic advective velocity of a 'infinitesimal' volume of fluid. this volume is small enough to say infinitesimal, but still many time larger than molecular scale. momentum density $\rho u$ is advected by $u$ gives rise to momentum flux $(\rho u)u$. what about the second term ?

p is pressure. it is really the coarsed-grained momentum flux of random motion of fluid molecules.

another way of thinking about pressure is to think about momentum conservation of a control volume of fluid at rest. consider in 1D, volume [a,b]. at the left boundary x=a, and non-zero temperature, the molecules have random motions. the molecules moving right at x=a will add positive momentum flux through left boundary, and the molecules moving left at x=a also add positive momentum flux (because they have negative momentum and they are leaving the volume). so p is here to take care of the microscopic momentum flux through the left boundary, and it is always positive. of course at the right boundary, if p is the same, then the net microscopic momentum flux is zero. if p is different at the right boundary, then there is a net momentum flux, which will gives rise to a macroscopic motion of the fluid.

heuristically, p = F/A (force / unit area) = (dP/dt)/A (rate of change of momentum per unit area). this is precisely an expression of momentum flux: momentum per unit area per unit time.

in 1D the conservation of momentum in derivative form is:

$$\frac{\partial(\rho u)}{\partial t} = \frac{\partial(\rho u^2 + p)}{\partial x}$$

this is the usual form of conservation law of some quantity: its time derivative = its flux. the role of p is: gradient of microscopic momentum flux gives right to macroscopic change in momentum (LHS). That is pretty cool.

viscous drag is also momentum flux. but this momentum is not normal to the surface but parallel to the surface. more on that later.