CNS 187 - Neural Computation
Solution Set 0

It is highly recommended that everyone download and work through the linear algebra review, now available on the class web site.

0.1 Linear Algebra

- For the two generalized matrices

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \tag{1}$$

a) Write down $\mathbf{AB}$.

$$\mathbf{AB} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \tag{2}$$

b) For general $\mathbf{A}$ of size $n \times l$ and $\mathbf{B}$ of size $l \times m$, write the component $\mathbf{AB}_{ik}$ in terms of a sum over the products of the elements of $\mathbf{A}$, a_{ij} and $\mathbf{B}$, b_{jk}.

$$\mathbf{AB}_{ik} = \sum_{j=1}^{l} a_{ij} b_{jk} \tag{3}$$

- If $\mathbf{A} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

c) determine the eigenvalues and eigenvectors of $\mathbf{A}$

Given an n x n matrix $\mathbf{A}$, a scalar λ is an eigenvalue of $\mathbf{A}$ if there is a nonzero column vector $\mathbf{x}$ such that $\mathbf{Ax} = \lambda\mathbf{x}$.

This can be rewritten as $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, which has a non-trivial solution only when

$$det(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0} \tag{4}$$

$$\begin{vmatrix} -\lambda & -1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = -\lambda(\lambda^2 - 1) + (-\lambda - 1) + (1 + \lambda) = 0 \tag{5}$$

Solving for lambda, we find three eigenvalues, $\lambda = -1, 0, 1$.

To find the eigenvectors, solve the equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ for each eigenvalue.

$$\begin{pmatrix} -\lambda & -1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{6}$$

$$\begin{aligned} -\lambda x_1 - x_2 + x_3 &= 0 \\ x_1 - \lambda x_2 + x_3 &= 0 \\ x_1 + x_2 - \lambda x_3 &= 0 \end{aligned} \tag{7}$$

Solving the equations, we find that for

$\lambda = -1 \quad \mathbf{x} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}$

$\lambda = 0 \quad \mathbf{x} = \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}$

$\lambda = 1 \quad \mathbf{x} = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

or some constant multiple thereof.

d) Adding any $N\mathbf{I}$ to the matrix results in changing the eigenvalues, but not the eigenvectors. For the matrix $[\mathbf{A} + 2\mathbf{I}]$, the new eigenvalues are 1, 2, 3. Each eigenvalue has increased by two, but still corresponds to the same eigenvector.

0.2 Random Walk

It's 2am and the Rathskeller has just closed for the night. N drunk graduate students and professors pile out of the bar onto the olive walk. At every step, each stumbles one step of length Δ to the left or right along the sidewalk with equal probability. After R steps, what is

a) the mean position of the members of the group?

Take s_i^r to be the r^{th} step that person i takes. Each s_i^r is Δ with probability 1/2 and $-\Delta$ with probability 1/2. All the s_i^r are independent identically distributed (i.i.d.) random variables. Take $x_i(R)$ as the position of the i^{th} person as a function of the number of steps, R; $x_i(R)$ is also a random variable.

Then:

$$x_i(R) = \sum_{r=1}^{R} s_i^r \tag{8}$$

And the mean position of the group for a given trial, $m(R)$ is:

$$m(R) = \frac{1}{N} \sum_{i=1}^{N} x_i(R) \tag{9}$$

And the *expected* mean position of the group is:

$$\langle (m(R)) \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle x_i(R) \rangle = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N} \sum_{r=1}^{R} (\frac{1}{2}(-\Delta)) + \frac{1}{2}(\Delta)) = 0. \tag{10}$$

The expected position of any given individual is also 0, of course.

b) their r.m.s. distance from the bar?

The mean, however, can deviate from its expected value on any given trail, the amount of this deviation is also a random variable, with the value

$$\delta(R) = m(R) - \langle m(R) \rangle \tag{11}$$

The variance is then:

$$\sigma^2 = \langle \delta^2(R) \rangle = \left\langle (m(R) - \langle (m(R)) \rangle)^2 \right\rangle = \langle m(R)^2 \rangle \tag{12}$$

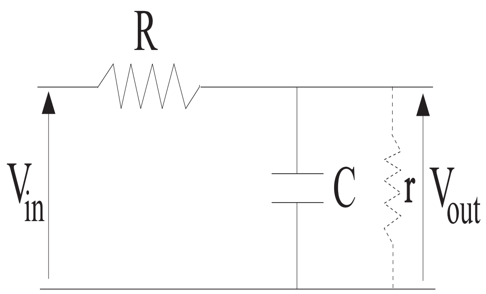
Because (1) variance is additive for i.i.d. variables, and we know that (2) multiplying a variable by a scalar multiplies its variance by the square of that scalar, we have:

$$\left\langle (m(R))^2 \right\rangle = \frac{1}{N^2} NR \left\langle (s_1^1 - \langle s_1^1 \rangle)^2 \right\rangle = \frac{R}{N} \left\langle (s_1^1)^2 \right\rangle = \frac{R}{N} \left(\frac{1}{2}(-\Delta)^2 + \frac{1}{2}(\Delta)^2 \right) = \frac{R\Delta^2}{N}. \tag{13}$$

Thus the r.m.s. distance of the group from the bar is $\frac{\sqrt{R}\Delta}{\sqrt{N}}$. Note that the r.m.s. distance of each individual is much greater: it is $\sqrt{R}\Delta$. The group (that is, their mean position) varies less than an given individual, on average.

0.3 RC Circuits

Consider the low-pass filter:



$$V_{in} = \begin{cases} 0V, & t < 0 \\ 5V, & t > 0 \end{cases} \tag{14}$$

a) Sketch the voltage response of the circuit as a function of time, with and without the parallel resistance, r.
b) Give the functional form of this response, again with and without the parallel resistance.

We can start with Kirchoff's Current Law (the sum of the currents out of a node equals zero) to set up the differential equation of the system as follows.

$$C\frac{dV_{out}}{dt} + \frac{V_{out} - V_{in}}{R} = 0 \tag{15}$$

Which we will rewrite as

$$\dot{V}_{out} + \frac{V_{out}}{RC} - \frac{V_{in}}{RC} = 0 \tag{16}$$

For $t < 0$, $\frac{dV_{out}}{dt} = 0$, and $V_{out} = 0$

For $t > 0$, assume a solution of the form $V_{out} = ae^{\lambda t} + \phi$. Since at t=0, $V_{out} = 0, \phi = -a$. Substituting this back into equation 16,

$$\lambda ae^{\lambda t} + \frac{ae^{\lambda t}}{RC} - \frac{a}{RC} - \frac{V_{in}}{RC} = 0 \tag{17}$$

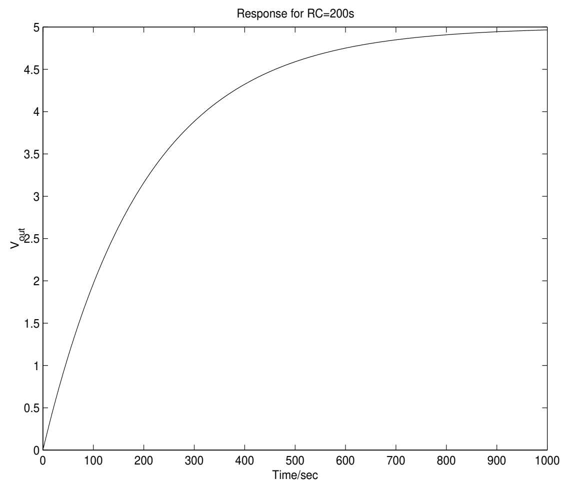
This has to be true for all $t > 0$. So it must be true that

$$\lambda a + \frac{a}{RC} = 0 \quad and \quad \frac{a}{RC} + \frac{5}{RC} = 0 \tag{18}$$

Therefore $\lambda = \frac{-1}{RC}$ and $a = -5$, giving us an equation for V_{out}

$$V_{out}(t) = 5(1 - e^{-\frac{t}{RC}}) \tag{19}$$

So the response function has the form



In the second case, we again use Kirchoff's Law to obtain the differential equation

$$\frac{V_{in} - V_{out}}{R} = C\frac{dV_{out}}{dt} + \frac{V_{out}}{r} \tag{20}$$

We can use the substitution as above, or more mechanistically use the Integrating Factor Method[1], for solving equations of this form.

$$-V_{out}e^{\int \frac{1}{C}(\frac{1}{r}+\frac{1}{R}).dt} = \int e^{\int \frac{1}{C}(\frac{1}{r}+\frac{1}{R}).dt}\frac{V_{in}}{RC}.dt \tag{21}$$

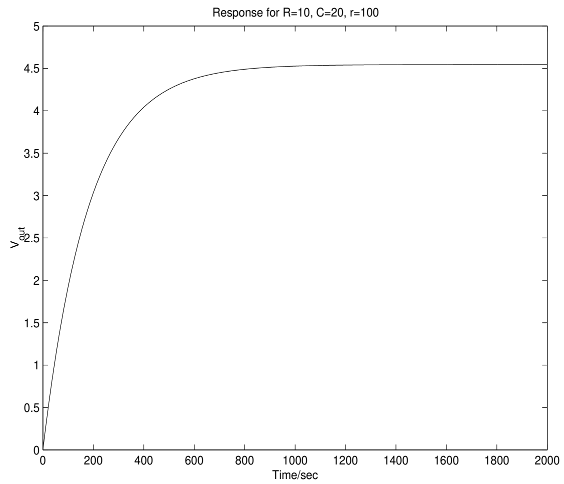
so

$$V_{out} = \frac{V_{in}}{RC}\left[\frac{C}{\frac{1}{r}+\frac{1}{R}} - De^{-\frac{t}{C}(\frac{1}{r}+\frac{1}{R})}\right] \tag{22}$$

At t=0, the capacitor is uncharged, so $V_{out} = 0$, hence we have

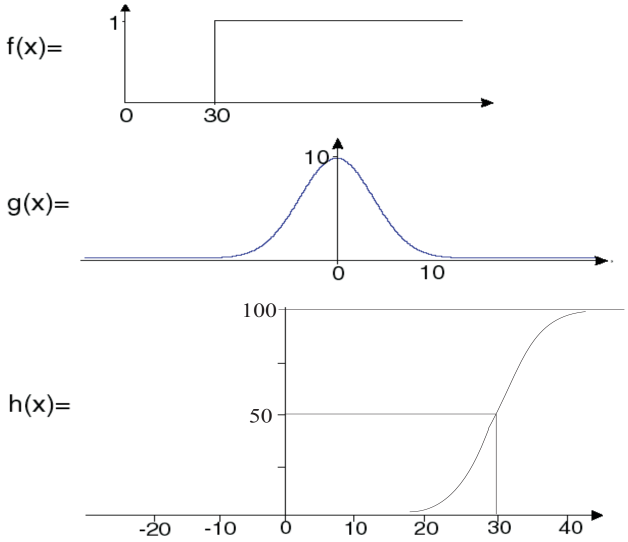
$$V_{out} = \frac{Vin}{\frac{R}{r}+1}\left[1 - e^{-\frac{t}{C}(\frac{1}{r}+\frac{1}{R})}\right] \tag{23}$$

which looks very similar to above, but has a shorter time constant, and V_{out} has a maximum value of $\frac{V_{in}}{\frac{R}{r}+1}$.



[1]This method is a completely general form for first order linear differential equations. See any diff-eq book, e.g. M. Braun, *Differential Equations and Their Applications*, pages 1–9.

0.4 Convolution



Given the form of the functions $f(x)$ and $g(x)$ in Figure 2, sketch roughly the form of

$$h(x) = \int_{-\infty}^{+\infty} g(s) f(x-s) ds \tag{24}$$

on the outline above, indicating the relative amplitudes at $x = 30$ and $x = +\infty$.

It is easy to do this problem if you realise two key points

- The value of the integral at each value of x is the area under the product f(x) and g(-s) aligned such that $s = 0$ is at x.
- $g(x)$ can be treated as roughly triangular to get the idea of what is going on at $x = 30$ and $x = +\infty$.

At $x = 30$ the area under the products is half the area of g(x), which is roughly 50. At $x \gg 30$ the product of the two functions is just $g(x)$ so $h(x)$ is equal in value to the area under $g(x)$, roughly 100.

0.5 Steepest Descent

$$q(x, y, z, s, t) = \sin(x + y) + z + \frac{1}{2s + 4t} \tag{25}$$

If you are at $x = 1, y = 1, z = 1, s = 1, t = 1$, and you take a very small step of length L in some direction, in what direction should you take the step to decrease q the most?

The direction of steepest descent of the function is $-\nabla q(x, y, z, s, t)$

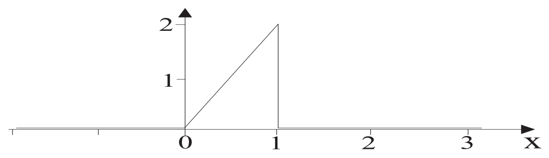
$$-\nabla q = \left(\frac{\partial q}{\partial x}\hat{x} + \frac{\partial q}{\partial y}\hat{y} + \frac{\partial q}{\partial z}\hat{z} + \frac{\partial q}{\partial s}\hat{s} + \frac{\partial q}{\partial t}\hat{t}\right) \quad (26)$$

$$= -cos(x+y)\hat{x} - cos(x+y)\hat{y} - \hat{z} + \frac{2}{(2s+4t)^2}\hat{s} + \frac{4}{(2s+4t)^2}\hat{t} \quad (27)$$

or

$$-\nabla q = \begin{pmatrix} -cos(x+y) \\ -cos(x+y) \\ -1 \\ \frac{2}{(2s+4t)^2} \\ \frac{4}{(2s+4t)^2} \end{pmatrix} = \begin{pmatrix} -cos(2) \\ -cos(2) \\ -1 \\ \frac{1}{18} \\ \frac{1}{9} \end{pmatrix} \quad (28)$$

0.6 Probability Density Functions

$p(x) =$ 

For the normalized probability density function, $p(x)$, write an integral expression in terms of $p(x)$ for:

a) the mean, or expectation value, $\langle x \rangle$.

$$\langle x \rangle = \int_{-\infty}^{+\infty} x p(x) dx \quad (29)$$

b) the variance, $\langle x^2 \rangle - \langle x \rangle^2$.

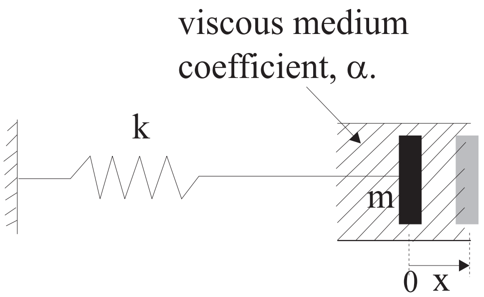
$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 p(x) dx \quad (30)$$

c) If $p(x)$ has the form shown in Figure 3, evaluate the mean and variance of x.

$$\begin{aligned} \langle x \rangle &= \int_0^1 2x^2 dx = \frac{2}{3}x^3 \mid_0^1 = \frac{2}{3} \\ \langle x^2 \rangle &= \int_0^1 2x^3 dx = \frac{1}{2}x^4 \mid_0^1 = \frac{1}{2} \\ \langle x^2 \rangle - \langle x \rangle^2 &= \frac{1}{2} - \frac{2^2}{3^2} = \frac{1}{18} \end{aligned} \quad (31)$$

0.7 Simple Oscillating Systems

A horizontal, damped mass-on-spring oscillator is shown in Figure 4.



If the viscous damping force is given by $-\alpha\dot{x}$, the equation of motion of the system is given by

$$m\ddot{x} + \alpha\dot{x} + kx = 0, \tag{32}$$

a) This may be written as

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} \text{ , where} \tag{33}$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -k/m & -\alpha/m \end{pmatrix}. \tag{34}$$

b) Recall that Eqn. 33 is solved by an expression of the form $\mathrm{e}^{\mathbf{A}t}\mathbf{y}(0)$, where $\mathrm{e}^{\mathbf{A}t} \equiv I + \mathbf{A}t + \frac{(\mathbf{A}t)^2}{2} + \frac{(\mathbf{A}t)^3}{3} + \cdots$ (if you'd like, plug this in to remind yourself why!). A convenient way to compute $\mathrm{e}^{\mathbf{A}t}$ in terms of the eigenvalues and eigenvectors of $\mathbf{A}$ is

$$\mathrm{e}^{\mathbf{A}t} = \mathbf{P}\mathrm{e}^{\mathbf{\Lambda t}}\mathbf{P}^{-1} \tag{35}$$

where (assuming that $\mathbf{A}$ is diagonalizable) $\mathbf{\Lambda}$ is the matrix

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{36}$$

of eigenvalues of $\mathbf{A}$, and $\mathbf{P}$ is the matrix whose columns are the corresponding eigenvectors $\mathbf{v1}$, $\mathbf{v2}$.

The solution to (33) is $\mathrm{e}^{\mathbf{A}t} = \mathbf{P}\mathrm{e}^{\mathbf{\Lambda}t}\mathbf{P}^{-1}\mathbf{y}(0)$ for arbitrary initial condition $\mathbf{y}(0) = \begin{pmatrix} y(0)_1 \\ y(0)_2 \end{pmatrix}$. Letting $\mathbf{P}^{-1}\mathbf{y}(0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, we have

$$\mathrm{e}^{\mathbf{A}t}\mathbf{y}(0) = c_1\mathrm{e}^{\lambda_1 t}\mathbf{v_1} + c_2\mathrm{e}^{\lambda_2 t}\mathbf{v_2}. \tag{37}$$

This shows how the eigenvalues and eigenvectors of $\mathbf{A}$ are related to the solution of (32).

The eigenvalues of $\mathbf{A}$ are $\lambda_{1,2} = 1/(2m)\left(-\alpha \pm \sqrt{\alpha^2 - 4km}\right)$, hence
$\mathbf{y}(t) = c_1 \exp\left[\frac{t}{2m}\left(-\alpha + \sqrt{\alpha^2 - 4km}\right)\right]\mathbf{v_1} + c_2 \exp\left[\frac{t}{2m}\left(-\alpha - \sqrt{\alpha^2 - 4km}\right)\right]\mathbf{v_2}$.

c) Note that for $k > k_{crit} \equiv \frac{\alpha^2}{4m}$, the arguments of both exponentials are imaginary, leading to oscillations of solutions around the equilibrium $x = v = 0$ (to see this, recall Euler's formula) (plot II). These oscillations decay, since the real part of these arguments is always negative. For $k < k_{crit}$, solutions simply decay to equilibrium (plot I).

The moral of this exercise is that qualitative information about the behavior of solutions to linear differential equation systems (or, locally, linearized nonlinear systems) of arbitrarily high dimension can often be gained by writing the equation in "matrix form" $\dot{y} = Ay$ and finding the eigenvalues of A.