CNS 187 Neural Computation - Problem Sheet 0 (Math Background)

Ling Li, ling@cs.caltech.edu

October 2, 2000

0.1 Linear Algebra

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

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

c) The eigenvalues of $\mathbf{A}$ are -1, 0, and 1, since the characteristic polynomial of $\mathbf{A}$ is $|\mathbf{A} - \lambda\mathbf{I}| = \lambda(1-\lambda)(1+\lambda)$. The corresponding eigenvectors (directions) are

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

d) Say, λ is one eigenvalue of $\mathbf{A}$ and x is one corresponding eigenvector. Then

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = |\mathbf{A} + 2\mathbf{I} - (\lambda + 2)\mathbf{I}|$$

and

$$(\mathbf{A} + 2\mathbf{I})\, x = (\lambda + 2)x.$$

So, after the diagonal values changed to 2, the eigenvalues of $\mathbf{A}$ should be 1, 2, and 3, and the corresponding eigenvectors remain unchanged.

0.2 Random Walk

Suppose N sufficiently large, then (a) 0; (b) $\sqrt{r}\Delta$

Define i.i.d. random variable $X \in \{+\Delta, -\Delta\}$ as the displacement of each step, and $x_{n,i}$ is the displacement one student or professor n $(= 1..N)$ makes at step i. Thus after r steps, the position of person n is $\sum_{i=1}^{r} x_{n,i}$, and the mean position and r.m.s. distance are

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{r} x_{n,i}, \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\sum_{i=1}^{r} x_{n,i}\right)^2}.$$

For N sufficiently large,

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{r}x_{n,i} = rEX = 0$$

since $+\Delta$ and $-\Delta$ have equal probability. And,

$$\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\sum_{i=1}^{r}x_{n,i}\right)^2} = \sqrt{r\cdot E(X^2)} = \sqrt{r}\Delta$$

since $x_{n,i}$ is independent with $x_{n,j}$ for those $i \neq j$.

For the case that r is sufficiently large but N is not so large, I guess that answer is (a) 0; (b) 0.

0.3 RC Circuits

From the circuit we have

$$C\frac{dV_{\text{out}}}{dt} = \frac{V_{\text{in}} - V_{\text{out}}}{R} - \frac{V_{\text{out}}}{r}.$$

Since R, C and r are constants, the equation can be solved as

$$V_{\text{out}} = e^{-\frac{R+r}{RCr}t}\left(\int_0^t \frac{V_{\text{in}}}{RC}e^{\frac{R+r}{RCr}x}dx + V_{\text{out}}|_{t=0}\right).$$

This may seem terrible. Fortunately, V_{in} doesn't change frequently. For $t < 0$, $V_{\text{in}} = 0$, we get

$$V_{\text{out}} = e^{-\frac{R+r}{RCr}t}\,V_{\text{out}}|_{t=0}.$$

Since physically V_{out} cannot be infinity when $t \to -\infty$, we can conclude $V_{\text{out}}|_{t=0}$ is 0. Thus for $t < 0$, $V_{\text{out}} = 0$. When $t > 0$, $V_{\text{in}} = 5(V)$,

$$\begin{aligned} V_{\text{out}} &= e^{-\frac{R+r}{RCr}t}\cdot\int_0^t \frac{5}{RC}e^{\frac{R+r}{RCr}x}dx \\ &= \frac{5r}{R+r}\left(1 - e^{-\frac{R+r}{RCr}t}\right). \end{aligned}$$

So, combining results above, we get

$$V_{\text{out}} = \begin{cases} 0, & t \leq 0 \\ \frac{5r}{R+r}\left(1 - e^{-\frac{R+r}{RCr}t}\right), & t > 0 \end{cases} \tag{1}$$

For the case without the parallel resistance r, we can substitute r in (1) with ∞, i.e.,

$$V_{\text{out}} = \begin{cases} 0, & t \leq 0 \\ 5\left(1 - e^{-\frac{1}{RC}t}\right), & t > 0 \end{cases}$$

The responses with and without r are sketched in Figure 1.[1]

[1] Illustrated by Matlab with $R = 100\Omega$, $C = 0.02F$, $r = 150\Omega$ or $r = \infty$.

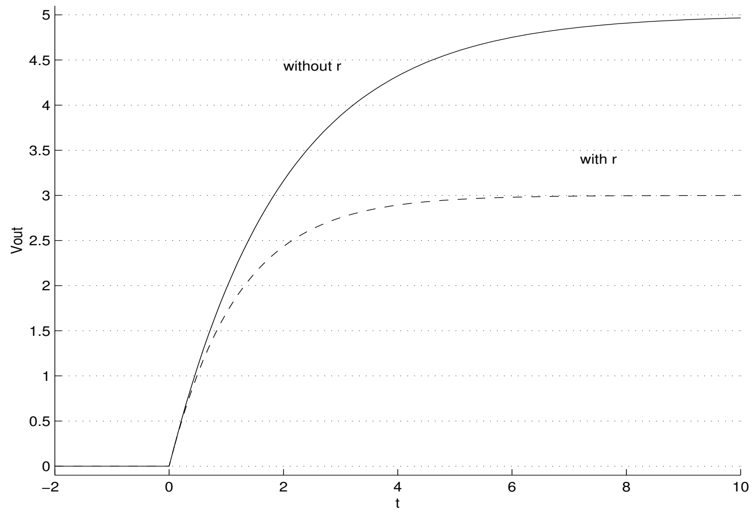

Figure 1: V_{out} with and without r.

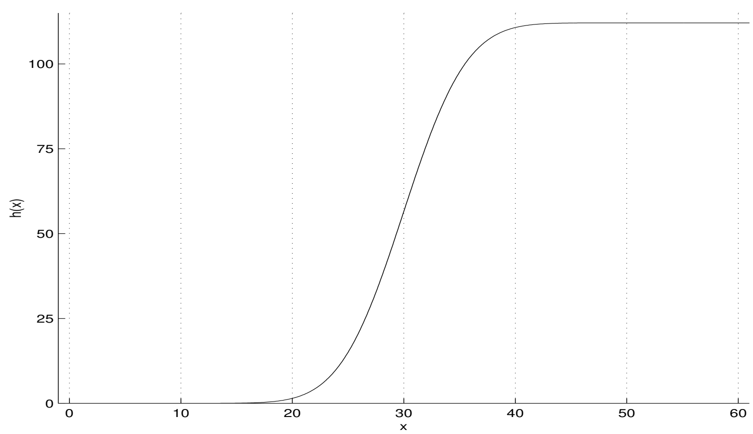

Figure 2: Convolution of $f(x)$ and $g(x)$.

0.4 Convolution

$$f(x) = \begin{cases} 1, & x > 30 \\ 0, & x \le 30 \end{cases}$$

$$h(x) = \int_{-\infty}^{+\infty} g(s) f(x-s) ds = \int_{-\infty}^{x-30} g(s) ds$$

$h(x)$ is roughly plotted in Figure 2, and we have[2]

$$h(0) \approx 0, h(+\infty) = 2h(30) \approx 112.$$

[2]We use $g(x)$ as $e^{-\frac{x^2}{40}}$ for illustration.

0.5 Steepest Descent

Study the partial differential of q against all its parameters at point $(1,1,1,1,1)$:

$$\begin{aligned}
\frac{\partial q}{\partial x} = \frac{\partial q}{\partial y} &= \cos(x+y) = \cos 2 \\
\frac{\partial q}{\partial z} &= 1 \\
\frac{\partial q}{\partial s} &= -\frac{2}{(2s+4t)^2} = -\frac{1}{18} \\
\frac{\partial q}{\partial t} &= -\frac{4}{(2s+4t)^2} = -\frac{1}{9}
\end{aligned}$$

Thus the steepest descent direction is

$$\Delta = -\left(\begin{array}{ccccc} \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} & \frac{\partial q}{\partial z} & \frac{\partial q}{\partial s} & \frac{\partial q}{\partial t} \end{array} \right)^T = \left(\begin{array}{c} -\cos 2 \\ -\cos 2 \\ -1 \\ \frac{1}{18} \\ \frac{1}{9} \end{array} \right).$$

0.6 Probability Density Functions

We may find $p(x)$ in the figure is

$$p(x) = \begin{cases} 2x, & 0 \le x \le 1; \\ 0, & \text{otherwise.} \end{cases}$$

a) the mean

$$\langle x \rangle = \int_{-\infty}^{+\infty} p(x) \cdot x dx = \int_0^1 2x^2 dx = \frac{2}{3}$$

b) the variance

$$\begin{aligned}
\left\langle x^2 \right\rangle - \langle x \rangle^2 &= \int_{-\infty}^{+\infty} p(x) \cdot x^2 dx - \left(\int_{-\infty}^{+\infty} p(x) \cdot x dx \right)^2 \\
&= \int_0^1 2x^3 dx - \left(\frac{2}{3} \right)^2 = \frac{1}{18}.
\end{aligned}$$

0.7 Simple Oscillating Systems

$$\begin{aligned}
\dot{x} &= v \\
\dot{v} &= \ddot{x} = -\frac{a}{m}\dot{x} - \frac{k}{m}x \\
&= -\frac{k}{m}x - \frac{a}{m}v
\end{aligned}$$

a) So

$$\begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{a}{m} \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix},$$

in the form of $\dot{\mathbf{y}} = \mathbf{A}\mathbf{y}$.
Suppose that $\mathbf{A}$ has two eigenvalues λ_1 and λ_2,[3] and two corresponding eigenvectors are $\mathbf{v}_1$ and $\mathbf{v}_2$. Define $\mathbf{V} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix}$, we have

$$\mathbf{A} \cdot \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{v}_1 & \lambda_2 \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

namely,

$$\mathbf{A} = \mathbf{V} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{V}^{-1}.$$

Under this situation, the solution to $\dot{\mathbf{y}} = \mathbf{A}\mathbf{y}$ is

$$\begin{aligned} \mathbf{y} &= e^{\mathbf{A}t} \cdot \mathbf{k} = \mathbf{V} \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \mathbf{V}^{-1}\mathbf{k} \\ &= \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= c_1 e^{\lambda_1 t} \mathbf{v}_1 + c_2 e^{\lambda_2 t} \mathbf{v}_2, \end{aligned}$$

where $\mathbf{k}$ is a 2×1 vector and $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathbf{V}^{-1}\mathbf{k}$.

b) Thus we know that answer for (b) is: λ_i and $\mathbf{v}_i$ are $\mathbf{A}$'s eigenvalues and eigenvectors, and

$$\lambda_{1,2} = \frac{-a \pm \sqrt{a^2 - 4mk}}{2m}.$$

c) When $a^2 - 4mk > 0$, i.e., $k < \dfrac{a^2}{4m}$, $\lambda_{1,2}$ are negtive real numbers. The displacement x decreases when time t increases. If considered physically, the spring elastic coefficient k is not big enough to make the object m oscillate in the viscous medium, so m will gradually get closer to its balance position. This is plot I.
When $a^2 - 4mk < 0$, i.e., $k > \dfrac{a^2}{4m}$, $\lambda_{1,2}$ are conjugate complex numbers. Under this situation,

$$e^{\lambda_{1,2} t} = e^{-\frac{a}{2m}t} \left[\cos\left(\frac{\sqrt{4mk - a^2}}{2m} t \right) \pm i \cdot \sin\left(\frac{\sqrt{4mk - a^2}}{2m} t \right) \right]$$

and displacement x will be

$$x = x(0) \cdot e^{-\frac{a}{2m}t} \cos\left(\frac{\sqrt{4mk - a^2}}{2m} t \right).$$

That is plot II. Physically, the spring elastic coefficient k is now so large that the spring can 'pull' or 'push' the object againt the viscous damping force to form a decreasing oscillating.

[3] λ_1 and λ_2 can be complex numbers.