# Problem 1

We would like to maximize

$$\left(\prod_{i=1}^{m}\sum_{z^{(i)}} p\left(x^{(i)}, z^{(i)}|\theta\right)\right) p(\theta)$$

Taking the log-likelihood we have

$$\begin{aligned}
\ell(\theta) &= \log p(\theta) + \sum_{i=1}^{m} \log \sum_{z^{(i)}} p\left(x^{(i)}, z^{(i)}|\theta\right) \\
&= \log p(\theta) + \sum_{i=1}^{m} \log \sum_{z^{(i)}} Q_i(z^{(i)}) \frac{p\left(x^{(i)}, z^{(i)}|\theta\right)}{Q_i(z^{(i)})} \\
&\geq \log p(\theta) + \sum_{i=1}^{m} \sum_{z^{(i)}} Q_i(z^{(i)}) \log \left[\frac{p\left(x^{(i)}, z^{(i)}|\theta\right)}{Q_i(z^{(i)})}\right]
\end{aligned}$$

where I have used Jensen's inequality in getting to the last line. Now the E-step is (as in the lecture notes) found by making this inequality tight. By this we set

$$\begin{aligned}
& Q_i(z^{(i)}) \propto p(x^{(i)}, z^{(i)}|\theta) \\
\implies \; & Q_i(z^{(i)}) = \frac{p(x^{(i)}, z^{(i)}|\theta)}{\sum_{z^{(i)}} p(x^{(i)}, z^{(i)}|\theta)} = p\left(z^{(i)}|x^{(i)};\theta\right)
\end{aligned}$$

In the $M$ step we must now maximize this lower bound:

$$\theta := \arg\max_{\theta} \left[\log p(\theta) + \sum_{i=1}^{m}\sum_{z^{(i)}} Q_i(z^{(i)}) \log \frac{p\left(x^{(i)}, z^{(i)}|\theta\right)}{Q_i(z^{(i)})}\right]$$

To show that this step is tractable, note that we are finding the maximum of a concave function (i.e. the sum of two concave functions is itself concave), and so there is a unique maximum.

The proof that $\left(\prod_{i=1}^{m}\sum_{z^{(i)}} p\left(x^{(i)}, z^{(i)}|\theta\right)\right) p(\theta)$ monotonically increases with each iteration is exactly as in the lecture notes. Each time we pick a $Q$ for some given $\theta^t$ in the $E$ step, we are picking a tight lower bound on the log-likelihood, $\ell(\theta^t)$. In the $M$ step, we improve (i.e. maximize) that lower bound by choosing an optimal $\theta^{t+1}$, and so we necessarily have that $\ell(\theta^{t+1}) \geq \ell(\theta^t)$.

# Problem 2

**(a)i** We begin with the assumption that the reviewer's scores are generated according to a random process given by

$$\begin{aligned} y^{(pr)} &\sim \mathcal{N}\left(\mu_p, \sigma_p^2\right), \\ z^{(pr)} &\sim \mathcal{N}\left(\nu_r, \tau_r^2\right), \\ \epsilon^{(pr)} &\sim \mathcal{N}\left(0, \sigma^2\right), \\ x^{(pr)} &= y^{(pr)} + z^{(pr)} + \epsilon^{(pr)} \end{aligned}$$

with $\epsilon, y^{(pr)}$ and $z^{(pr)}$ all independent. Our joint distribution will be Gaussian:

$$\begin{bmatrix} x^{(pr)} \\ y^{(pr)} \\ z^{(pr)} \end{bmatrix} \sim \mathcal{N}\left(\mu_{xyz}, \Sigma_{xyz}\right)$$

For the expectation value of $x^{(pr)}$ we use

$$\mathrm{E}[x^{(pr)}] = \mathrm{E}[y^{(pr)} + z^{(pr)} + \epsilon^{(pr)}] = \mathrm{E}[y^{(pr)}] + \mathrm{E}[z^{(pr)}] + \mathrm{E}[\epsilon^{(pr)}] = \mu_p + \nu_r$$

and so the mean vector is

$$\mu_{xyz} = \begin{bmatrix} \mu_p + \nu_r \\ \mu_p \\ \nu_r \end{bmatrix}$$

Let us now calculate each term in the covariance matrix $\Sigma_{xyz}$. The $yy$ term is necessarily $\sigma_p^2$ while the $zz$ term is necessarily $\tau_r^2$. Next, the $yz$ and $zy$ terms are necessarily zero because of our assumption that $y^{(pr)}$ and $z^{(pr)}$ are independent variables. The $xy$ term is given by

$$\begin{aligned} \mathrm{E}[(y^{(pr)} - \mu_p)(x^{(pr)} - \mu_p - \nu_r)] &= \mathrm{E}[(y^{(pr)} - \mu_p)(y^{(pr)} + z^{(pr)} + \epsilon^{(pr)} - \mu_p - \nu_r)] \\ &= \mathrm{E}[(y^{(pr)} - \mu_p)(y^{(pr)} - \mu_p) + (y^{(pr)} - \mu_p)(z^{(pr)} - \nu_r) + (y^{(pr)} - \mu_p)(\epsilon^{(pr)})] \\ &= \mathrm{E}[(y^{(pr)} - \mu_p)(y^{(pr)} - \mu_p)] + \mathrm{E}[(y^{(pr)} - \mu_p)(z^{(pr)} - \nu_r)] + \mathrm{E}[(y^{(pr)} - \mu_p)(\epsilon^{(pr)})] \\ &= \sigma_p^2 + 0 + 0 \end{aligned}$$

Similarly, we find that the $xz$ term is given by $\mathrm{E}[(z^{(pr)} - \nu_p)(x^{(pr)} - \mu_p - \nu_r)] = \tau_r^2$. Finally, for the $xx$ term we have

$$\begin{aligned} \mathrm{E}[(x^{(pr)} - \mu_p - \nu_r)(x^{(pr)} - \mu_p - \nu_r)] &= \mathrm{E}[(y^{(pr)} + z^{(pr)} + \epsilon^{(pr)} - \mu_p - \nu_r)(y^{(pr)} + z^{(pr)} + \epsilon^{(pr)} - \mu_p - \nu_r)] \\ &= \sigma_p^2 + \tau_r^2 + \sigma^2 \end{aligned}$$

where in getting to this second line we have used the fact that there are no non-zero covariances among these variables. Putting all these results together we find that the joint Gaussian distribution takes the form

$$\begin{bmatrix} x^{(pr)} \\ y^{(pr)} \\ z^{(pr)} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_p + \nu_r \\ \mu_p \\ \nu_r \end{bmatrix}, \begin{bmatrix} \sigma_p^2 + \tau_r^2 + \sigma^2 & \sigma_p^2 & \tau_r^2 \\ \sigma_p^2 & \sigma_p^2 & 0 \\ \tau_r^2 & 0 & \tau_r^2 \end{bmatrix}\right) \tag{1}$$

**(a)ii** The rules for conditioning on subsets of jointly Gaussian random variables are

$$\begin{aligned} \mu_{1|2} &= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) \\ \Sigma_{1|2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \end{aligned}$$

Applying these formulae we find

$$\mu_{Q_{pr}} = \begin{bmatrix} \mu_p + \frac{\sigma_p^2}{\sigma_p^2+\tau_r^2+\sigma^2}\left(x^{(pr)} - \mu_p - \nu_r\right) \\ \nu_r + \frac{\tau_r^2}{\sigma_p^2+\tau_r^2+\sigma^2}\left(x^{(pr)} - \mu_p - \nu_r\right) \end{bmatrix} \tag{2}$$

while the covariance matrix is given by

$$\begin{aligned}\Sigma_{Q_{pr}} &= \begin{bmatrix} \sigma_p^2 & 0 \\ 0 & \tau_r^2 \end{bmatrix} - \begin{bmatrix} \sigma_p^2 \\ \tau_r^2 \end{bmatrix} \frac{1}{\sigma_p^2 + \tau_r^2 + \sigma^2} \begin{bmatrix} \sigma_p^2 & \tau_r^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_p^2 & 0 \\ 0 & \tau_r^2 \end{bmatrix} - \frac{1}{\sigma_p^2 + \tau_r^2 + \sigma^2} \begin{bmatrix} \sigma_p^4 & \sigma_p^2\tau_r^2 \\ \sigma_p^2\tau_r^2 & \tau_r^4 \end{bmatrix} \\ &= \frac{1}{\sigma_p^2 + \tau_r^2 + \sigma^2} \begin{bmatrix} \sigma_p^2(\tau_r^2 + \sigma^2) & -\sigma_p^2\tau_r^2 \\ -\sigma_p^2\tau_r^2 & \tau_r^2(\sigma_p^2 + \sigma^2) \end{bmatrix} \end{aligned} \tag{3}$$

and so we find that $Q_{pr}(y^{(pr)}, z^{(pr)})$ is a Gaussian distribution,

$$Q_{pr}(y^{(pr)}, z^{(pr)}) \sim \mathcal{N}\left( \begin{bmatrix} \mu_p + \frac{\sigma_p^2}{\sigma_p^2+\tau_r^2+\sigma^2}\left(x^{(pr)} - \mu_p - \nu_r\right) \\ \nu_r + \frac{\tau_r^2}{\sigma_p^2+\tau_r^2+\sigma^2}\left(x^{(pr)} - \mu_p - \nu_r\right) \end{bmatrix}, \quad \frac{1}{\sigma_p^2 + \tau_r^2 + \sigma^2} \begin{bmatrix} \sigma_p^2(\tau_r^2 + \sigma^2) & -\sigma_p^2\tau_r^2 \\ -\sigma_p^2\tau_r^2 & \tau_r^2(\sigma_p^2 + \sigma^2) \end{bmatrix} \right)$$

**(b)** In the $M$ step we must optimize the function $f(\mu_p, \nu_r, \sigma_p^2, \sigma_r^2)$ where

$$f(\mu_p, \nu_r, \sigma_p^2, \sigma_r^2) = \sum_{p=1}^{P}\sum_{r=1}^{R} \int_{y^{(pr)}, z^{(pr)}} Q_{pr}(y^{(pr)}, z^{(pr)}) \log \frac{p(x^{(pr)}, y^{(pr)}, z^{(pr)}; \mu_p, \nu_r, \sigma_p^2, \tau_r^2)}{Q_{pr}(y^{(pr)}, z^{(pr)})}$$

where the $\int_{y^{(pr)}, z^{(pr)}}$ denotes some average of $y^{(pr)}$ and $z^{(pr)}$ variables. We note that as always in EM algorithms, $Q_{pr}$ contains 'old' values of $\mu_p, \nu_r \ldots$, and so when finding arg max of this function the only interesting dependence comes from the $Q_{pr} \log p(\cdots)$ term. Noting that we can write this term as an expectation value, i.e.

$$\begin{aligned}&\sum_{p=1}^{P}\sum_{r=1}^{R} \int_{y^{(pr)}, z^{(pr)}} Q_{pr}(y^{(pr)}, z^{(pr)}) \log p(x^{(pr)}, y^{(pr)}, z^{(pr)}; \mu_p, \nu_r, \sigma_p^2, \tau_r^2) \\ &\equiv \sum_{p=1}^{P}\sum_{r=1}^{R} \mathrm{E}_{\{y^{(pr)}, z^{(pr)}\} \sim Q_{pr}} \left[ \log p(x^{(pr)}, y^{(pr)}, z^{(pr)}; \mu_p, \nu_r, \sigma_p^2, \tau_r^2) \right]\end{aligned}$$

and I will henceforth shorten the notation for this expectation value to $\tilde{\mathrm{E}}$. Let us write out the form of $p(x^{(pr)}, y^{(pr)}, z^{(pr)}; \mu_p, \nu_r, \sigma_p^2, \tau_r^2)$ explicitly. Using the mean vector and covariance vector in Eq. 1 with a straightforward matrix inversion, we find

$$p(x^{(pr)}, y^{(pr)}, z^{(pr)}; \mu_p, \nu_r, \sigma_p^2, \tau_r^2) = \frac{1}{(2\pi)^{3/2}\sigma_p\tau_r\sigma} \exp\left[ -\frac{(y^{(pr)} - \mu_p)^2}{2\sigma_p^2} - \frac{(z^{(pr)} - \nu_r)^2}{2\tau_r^2} - \frac{(x^{(pr)} - y^{(pr)} - z^{(pr)})^2}{2\sigma^2} \right]$$

(This result is in fact exactly what we were given in the problem to begin with, so it serves as a useful consistency check!). Dropping terms that do not depend on the variables $\mu_p, \nu_r, \sigma_p^2$ and $\tau_r^2$ we have

$$f(\mu_p, \nu_r, \sigma_p^2, \sigma_r^2) = \sum_{p=1}^{P}\sum_{r=1}^{R} \tilde{\mathrm{E}} \left[ \log\left(\frac{1}{\sigma_p\tau_r}\right) - \frac{(y^{(pr)} - \mu_p)^2}{2\sigma_p^2} - \frac{(z^{(pr)} - \nu_r)^2}{2\tau_r^2} \right]$$

We can now expand out these terms and use the fact the following facts:

$$\tilde{\mathrm{E}}[(y^{(pr)})] = \mu^y_{Q_{pr}}, \quad \tilde{\mathrm{E}}[(z^{(pr)})] = \mu^z_{Q_{pr}} \quad \text{and}$$
$$\tilde{\mathrm{E}}[(y^{(pr)})^2] = \tilde{\mathrm{E}}[(y^{(pr)})^2] - \tilde{\mathrm{E}}[y^{(pr)}]^2 + \tilde{\mathrm{E}}[y^{(pr)}]^2 = \Sigma^{yy}_{Q_{pr}} + (\mu^y_{Q_{pr}})^2$$
$$\tilde{\mathrm{E}}[(z^{(pr)})^2] = \tilde{\mathrm{E}}[(z^{(pr)})^2] - \tilde{\mathrm{E}}[z^{(pr)}]^2 + \tilde{\mathrm{E}}[z^{(pr)}]^2 = \Sigma^{zz}_{Q_{pr}} + (\mu^z_{Q_{pr}})^2$$

where $\Sigma^{yy}_{Q_{pr}}$ is the $yy$ component of the matrix in Eq. 3, and $\mu^y_{Q_{pr}}$ is the $y$ component of the vector in Eq. 2 etc. Note that these would be numbers that were derived in the $E$ step. Using these definitions we find

$$f(\mu_p, \nu_r, \sigma_p^2, \sigma_r^2)$$
$$= \sum_{p=1}^{P}\sum_{r=1}^{R}\left[\log\left(\frac{1}{\sigma_p\tau_r}\right) - \frac{1}{2\sigma_p^2}\left(\Sigma^{yy}_{Q_{pr}} + (\mu^y_{Q_{pr}})^2 - 2\mu_p\mu^y_{Q_{pr}} + \mu_p^2\right) - \frac{1}{2\tau_r^2}\left(\Sigma^{zz}_{Q_{pr}} + (\mu^z_{Q_{pr}})^2 - 2\nu_r\mu^z_{Q_{pr}} + \nu_r^2\right)\right]$$

Now let us maximize this expression with respect to each of the parameters.

$$\frac{\partial f}{\partial \mu_p} = \sum_{r=1}^{R}\frac{1}{\sigma_p^2}\left(-\mu^y_{Q_{pr}} + \mu_p\right) = 0$$
$$\implies \mu_p = \frac{1}{R}\sum_{r=1}^{R}\mu^y_{Q_{pr}} \tag{4}$$

Similarly, $\frac{\partial f}{\partial \nu_r} = 0$ implies

$$\nu_r = \frac{1}{P}\sum_{r=1}^{P}\mu^z_{Q_{pr}} \tag{5}$$

Next, for the variances we have

$$\frac{\partial f}{\partial \sigma_p} = \sum_{r=1}^{R}\left[-\frac{1}{\sigma_p} + \frac{1}{\sigma_p^3}\left(\Sigma^{yy}_{Q_{pr}} + (\mu^y_{Q_{pr}})^2 - 2\mu_p\mu^y_{Q_{pr}} + \mu_p^2\right)\right] = 0$$
$$\implies \sigma_p^2 = \frac{1}{R}\sum_{r=1}^{R}\left(\Sigma^{yy}_{Q_{pr}} + (\mu^y_{Q_{pr}})^2 - 2\mu_p\mu^y_{Q_{pr}} + \mu_p^2\right) \tag{6}$$

and finally, a similar set of manipulations for $\tau_r^2$ gives

$$\tau_r^2 = \frac{1}{P}\sum_{p=1}^{P}\left(\Sigma^{zz}_{Q_{pr}} + (\mu^z_{Q_{pr}})^2 - 2\nu_r\mu^z_{Q_{pr}} + \nu_r^2\right) \tag{7}$$

So to summarize, in the E-step we must compute Equations 2 and 3, then in the M-step we calculate equations 4 through 7.

# Problem 3

We are given that

$$f_u(x) = \arg\min_{v \in \mathcal{V}} ||x - v||^2.$$

Another way of writing $||x - v||^2$ is

$$||x - v||^2 = (\vec{x} - \vec{v})^2 = \vec{x} \cdot \vec{x} + \vec{v} \cdot \vec{v} - 2\vec{x} \cdot \vec{v}$$

where I have denoted $x$ and $v$ has vectors in Euclidean space, with $\vec{x} \cdot \vec{v} = ||x||\,||v|| \cos\theta$ and $\theta$ is the angle between these two vectors. Now, writing $\vec{v} = \alpha\hat{u}$, where $\hat{u}$ is a unit length vector, we find

$$||x - v||^2 = \vec{x} \cdot \vec{x} + \alpha^2 - 2\alpha\vec{x} \cdot \hat{u}$$

Minimizing w.r.t to $\alpha$ we get that this is minimized when $\alpha = \alpha^*$ with

$$\alpha^* = \vec{x} \cdot \hat{u}$$

It then clearly follows that

$$f_u(x) \equiv \arg\min_{v \in \mathcal{V}} ||x - v||^2 = \alpha^* \hat{u} = (\vec{x} \cdot \hat{u})\hat{u}$$

Now substitute this into the given expression:

$$\begin{aligned}
\arg\min_{u:\, u^T u = 1} \sum_{i=1}^{m} ||x^{(i)} - f_u(x^{(i)})||_2^2 &= \arg\min_{u:\, u^T u=1} \sum_{i=1}^{m} \left(\vec{x}^{(i)} - (\vec{x}^{(i)} \cdot \hat{u})\hat{u}\right)^2 \\
&= \arg\min_{u:\, u^T u=1} \sum_{i=1}^{m} \left(\vec{x}^{(i)} \cdot \vec{x}^{(i)} - 2(\vec{x}^{(i)} \cdot \hat{u})^2 + (\vec{x}^{(i)} \cdot \hat{u})^2(\hat{u} \cdot \hat{u})\right) \\
&= \arg\min_{u:\, u^T u=1} \sum_{i=1}^{m} -(\vec{x}^{(i)} \cdot \hat{u})^2 \\
&= \arg\max_{u:\, u^T u=1} \sum_{i=1}^{m} (\vec{x}^{(i)} \cdot \hat{u})^2 \\
&= \arg\max_{u:\, u^T u=1} \sum_{i=1}^{m} \sum_{\alpha,\beta} u_\alpha \left(x_\alpha^{(i)} x_\beta^{(i)}\right) u_\beta \\
&= \arg\max_{u:\, u^T u=1} \sum_{\alpha,\beta} u_\alpha \left(\sum_{i=1}^{m} x_\alpha^{(i)} x_\beta^{(i)}\right) u_\beta
\end{aligned}$$

Where from the second to the third line we dropped the $\vec{x} \cdot \vec{x}$ term since it does not depend on $u$, and also used that fact that $\hat{u} \cdot \hat{u} \equiv u^T u = 1$, while in getting to the second to last line we have written the indices of the vectors explicitly, and in finding the last line we noticed that the sums can be interchanged. We then see that finding the maximum of this equation is akin to making $\hat{u}$ the largest eigenvector of the matrix $\Lambda$ whose components are $\Lambda_{\alpha\beta} = \left(\sum_{i=1}^{m} x_\alpha^{(i)} x_\beta^{(i)}\right)$, i.e. making $u$ the first principal component.

# Problem 4

The code for this problem is attached. The W matrix I found is shown in Fig. 1

W

5x5 double

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 64.7745 | 21.1275 | 25.8785 | -11.8901 | -13.5417 |
| 2 | 10.8774 | 22.4100 | -4.2617 | -15.7340 | 7.9671 |
| 3 | 17.8137 | -6.4520 | 28.9526 | 11.6666 | -15.1330 |
| 4 | -2.6242 | 0.5079 | -1.5239 | 6.7325 | -2.1902 |
| 5 | -5.9181 | 16.0100 | 11.2267 | 9.7081 | 27.8956 |

Figure 1: The W matrix found after running independent component analysis

# Problem 5

Begin by considering

$$B(V_1) - B(V_2) = R(s) - R(s) + \gamma \max_{a \in A} \sum_{s' \in S} P_{sa}(s')V_1(s') - \gamma \max_{a' \in A} \sum_{s'' \in S} P_{sa'}(s')V_2(s'')$$
$$\implies |B(V_1) - B(V_2)| = \gamma |\max_{a \in A} \sum_{s' \in S} P_{sa}(s')V_1(s') - \max_{a \in A} \sum_{s'' \in S} P_{sa'}(s')V_2(s'')|$$

Now to proceed let us prove that for two functions $W_1(a)$ and $W_2(a)$, $|\max_a W_1(a) - \max_a W_2(a)| \leq \max_a |W_1(a) - W_2(a)|$. First, assume $\arg\max_a W_1(a) = a_1^*$ and $\arg\max_a W_2(a) = a_2^*$. It is then generally the case that $|W_1(a_1^*) - W_2(a_2^*)| \geq 0$. It is also generally the case that

$$W_1(a_1^*) - W_2(a_2^*) \leq W_1(a_1^*) - W_2(a_1^*) \leq |W_1(a_1^*) - W_2(a_1^*)| \leq \max_a |W_1(a) - W_2(a)|$$

where we have used the fact $W_2(a_2^*) \geq W_2(a_1^*)$ in deriving the first inequality. It is also the case that

$$W_2(a_2^*) - W_1(a_1^*) \leq W_2(a_2^*) - W_1(a_2^*) \leq |W_2(a_2^*) - W_1(a_2^*)| \leq \max_a |W_2(a) - W_1(a)|$$

Both of these statements imply the statement:

$$|\max_a W_1(a) - \max_a W_2(a)| \leq \max_a |W_1(a) - W_2(a)|$$

So using this we can go from the first to the second line below:

$$\begin{aligned}
||B(V_1) - B(V_2)||_\infty &= \gamma \max_{s \in S} \left| \max_{a \in A} \sum_{s' \in S} P_{sa}(s')V_1(s') - \max_{a \in A} \sum_{s'' \in S} P_{sa'}(s')V_2(s'') \right| \\
&\leq \gamma \max_{s \in S} \max_{a \in A} \left| \sum_{s' \in S} P_{sa}(s')V_1(s') - \sum_{s'' \in S} P_{sa'}(s')V_2(s'') \right| \\
&= \gamma \max_{s \in S} \max_{a \in A} \left| \sum_{s' \in S} P_{sa}(s') \left(V_1(s') - V_2(s')\right) \right| \\
&\leq \gamma \max_{s \in S} \max_{a \in A} \sum_{s' \in S} P_{sa}(s') \big| V_1(s') - V_2(s') \big| \\
&\leq \gamma \max_{s \in S} \max_{a \in A} \sum_{s' \in S} P_{sa}(s') \big|\big| V_1 - V_2 \big|\big|_\infty \\
&= \gamma \max_{s \in S} \max_{a \in A} \big|\big| V_1 - V_2 \big|\big|_\infty \sum_{s' \in S} P_{sa}(s') \\
&= \gamma \big|\big| V_1 - V_2 \big|\big|_\infty \\
\implies ||B(V_1) - B(V_2)||_\infty &\leq \gamma \big|\big| V_1 - V_2 \big|\big|_\infty
\end{aligned}$$

where in going from the second to the third line we noticed that $s'$ and $s''$ are dummy indices (being summed over) and so can be combined. Next in going from the third to the fourth line we use a triangle inequality. From the fourth to the fifth line we notice that the sum over elements is less than a sum where each element is replaced by the max norm (i.e. replace each term in the sum, the same element which is the largest element in the sum). For the sixth line to the 7th we have done the sum over $s'$ which gives one (definition of probabilities), and finally we removed the redundant max conditions.

**(b)** Let us imagine there is more than 1 fixed point. i.e. consider $B(V_1) = V_1$ and $B(V_2) = V_2$. From the above result we have

$$||V_1 - V_2||_\infty \leq \gamma ||V_1 - V_2||_\infty \implies (1 - \gamma)||V_1 - V_2||_\infty \leq 0 \tag{8}$$

This equation must be satisfied for all $\gamma$, and because both $||V_1 - V_2||_\infty$ and $(1 - \gamma)$ are always non-negative (and the latter is generally non-zero), this can only be true if $V_1 = V_2$, i.e. $B(V) = V$ only occurs for a single $V$.

# Problem 6

The code for this problem is attached. The typical number of trials before the pole falls off is in the range of 100 to 300. Shown below are two examples of learning curves. in the first, it takes 214 attempts before converging , while in the second it took 181.

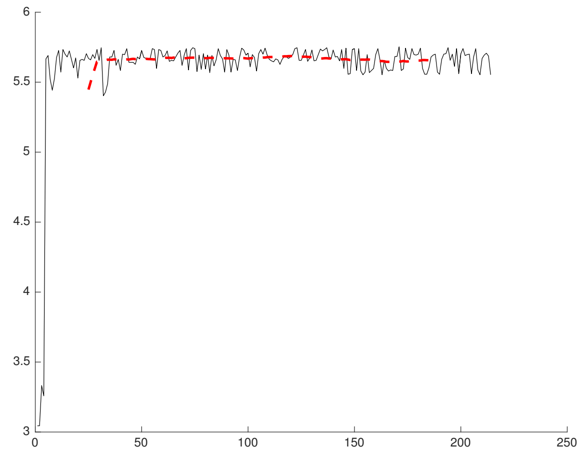


Figure 2: An example of the learning curve log(number of time steps before failure) vs. number of time failures. This took 214 attempts before convergence.

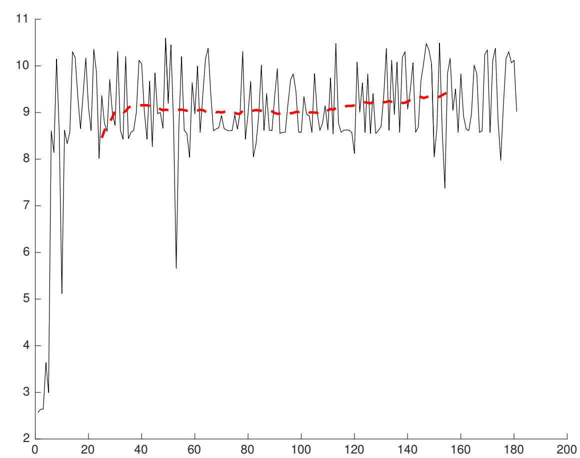


Figure 3: A second example of the learning curve log(number of time steps before failure) vs. number of time failures. This took 181 attempts before convergence, and appeared to have learned *really well!*. Certainly not a typical example!