Solutions to Review Problems

* Please contact me (soner@princeton.edu), if you have any comments or corrections.
* I will update and correct these notes continuously. Please make sure to have the latest version.

Version = 3, October 21 2025

Problem 1.1

a) Since $a^* = 1$, we only use the first time for all times after one, i.e., $a_k = 1$ for every $k \geq 2$. Hence the distribution of r_k is uniform of $[1,3]$ for all $k \geq 2$. Therefore,

$$J = \lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T} r_k = \lim_{T\to\infty} \frac{1}{T} \sum_{k=2}^{T} r_k$$

$$= \text{expected value of the rewards from the first arm}$$

$$= \int_1^3 \frac{1}{2} r \, dr = \boxed{2}$$

b) Same reasoning as in part a) implies that

$$J = \text{expected value of the rewards from the second arm}$$

$$= \int_2^4 \frac{1}{2} r \, dr = \boxed{3}$$

c) As $N=1$, we use both arms once. Let R_1 be the reward from arm one and R_2 be the one from the second.

Then, $\hat{q}(a) = R_a$ for $a=1,2$, and $a^*=1$ iff $R_1 > R_2$

Hence, $\mathbb{P}(a^*=1) = \mathbb{P}(R_1 \geq R_2)$. As R_1, R_2 are independent,

$$f_{R_1,R_2}(x,y) = f_{R_1}(x) f_{R_2}(y), \text{ and}$$

$$f_{R_1}(x) = \frac{1}{2} \text{ for } x \in [1,3], \text{ zero otherwise},$$

$$f_{R_2}(y) = \frac{1}{2} \text{ for } y \in [2,4], \text{ zero otherwise.}$$

Therefore,

$$\mathbb{P}(R_1 > R_2) = \frac{1}{4} \int_2^4 \int_1^3 \chi_{\{x \geq y\}} \, dx \, dy$$

where

$$\chi_{\{x > y\}} = \begin{cases} 1 & , \text{ if } x \geq y, \\ 0 & , \text{ if } x < y. \end{cases}$$

So

$$\mathbb{P}(a^* = 1) = \frac{1}{4} \int_2^3 \int_y^3 dx \, dy = \frac{1}{4} \int_2^3 (3-y) \, dy = \frac{1}{8}.$$

Summarizing,

$$\boxed{\mathbb{P}(a^* = 1) = \frac{1}{8} \quad , \quad \mathbb{P}(a^* = 2) = \frac{7}{8}}$$

d)

$$\mathbb{E}[J] = \mathbb{E}[J \mid a^* = 1] \, \mathbb{P}(a^* = 1) + \mathbb{E}[J \mid a^* = 2]$$

$$= 2 \cdot \frac{1}{8} + 3 \cdot \frac{7}{8} = \boxed{\frac{23}{8}}$$

e) As for large k, the maximizer will be the second arm, on average we will chose the second arm $(90+5)\%$ of the time, and the first arm 5% of the time. Hence,

$$\text{limiting } J \text{ value} = (0.95) \, \mathbb{E}[r_t \mid a=2] + (0.05) \, \mathbb{E}[r_t \mid a=1]$$

$$= (0.95) 3 + (0.05) 2 = \boxed{2.95}$$

See the solutions to homework 1 for a full proof, but that is not required here.

Problem 1.2 Here we give only the computations as the reasoning is exactly as in Problem 1. Reward distributions:

arm 1: reward $=0$ or 1 with probability $\frac{1}{2}$ each;

arm 2: reward $=0$ with pr. $\frac{3}{4}$ and otherwise 1.

a) $J = \mathbb{E}[\text{reward} \mid \text{arm}=1] = \boxed{\frac{1}{2}}$

b) $J = \mathbb{E}[\text{reward} \mid \text{arm}=2] = \boxed{\frac{1}{4}}$

c) THIS IS SOLUTION WITH N=1

$$\mathbb{P}(a^*=2) = \mathbb{P}(R_1 < R_2) = \mathbb{P}(R_1=0 \text{ and } R_2=1) = \frac{1}{2}\cdot\frac{1}{4} = \boxed{\frac{1}{8}}$$

$$\Rightarrow \mathbb{P}(a^*=1) = \boxed{\frac{7}{8}}.$$

d) $\mathbb{E}[J] = \frac{1}{2}\cdot\frac{7}{8} + \frac{1}{4}\cdot\frac{1}{8} = \boxed{\frac{15}{32}}.$

e) We use first arm 90% and second arm 10%.

$$\text{limiting } J = (0.9)\frac{1}{2} + (0.1)\frac{1}{4} = \frac{1.9}{4} = \boxed{0.475}.$$

Problem 2.1

a) The controbility matrix is $\mathcal{C} = [B \,\vdots\, AB]$ and

$$AB = \begin{bmatrix} 1/2 & 0 \\ 0 & a \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}.$$

$\Rightarrow \mathcal{C} = \begin{bmatrix} 0 & 0 \\ 1 & a \end{bmatrix}$ which has rank $1<2$ $\Rightarrow$ not controllable $\forall a$.

b) Set $x_k = (v_k, s_k)$. Then,

$$v_{k+1} = \frac{1}{2} v_k, \quad \forall k = 1, 2 \ldots$$

$$s_{k+1} = a s_k + u_k, \quad \forall k = 1, 2 \ldots$$

If we chose $u_1 = -a s_0$, then $s_k = 0$ for all $k \geq 1$. But $v_k = 2^{-k} v_0$ for all $k \geq 1$ no matter which control we use.

Therefore, $V < \infty$ if one of the followings hold

either $v_0 = 0$ (then any $\rho > 0$, a and s_0 allowed)

or $v_0 \neq 0$, but $\rho < 4$ (any a and s_0 allowed).

c) $V = M + \rho A^T V A - \rho^2 A^T V B (N + \rho B^T V B)^{-1} B^T V A$.

Let $V = \begin{bmatrix} x & y \\ y & z \end{bmatrix}$. We first calculate that

$$N + B^T V B = 1 + z$$

Also,

$$A^T V B = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & y \\ y & z \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} y/2 \\ z \end{bmatrix}$$

$$B^T V A = (A^T V B)^T = [y/2, z]$$

and

$$A^T V A = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & y \\ y & z \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x/2 & y \\ y/2 & z \end{bmatrix} = \begin{bmatrix} x/4 & y/2 \\ y/2 & z \end{bmatrix}.$$

Therefore,

$$\boxed{\begin{bmatrix} x & y \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} x/4 & y/2 \\ y/2 & z \end{bmatrix} - \frac{1}{1+z} \begin{bmatrix} y^2/4 & yz/2 \\ yz/2 & z^2 \end{bmatrix}.}$$

d) We have 3 equations for 3 unknowns x, y, z:

$$x = 1 + \frac{1}{4}x - \frac{1}{4(1+z)}y^2 \qquad (1)$$

$$y = \frac{1}{2}y - \frac{1}{2(1+z)}yz \qquad (2)$$

$$z = 1 + z - \frac{1}{1+z}z^2 \qquad (3)$$

Third equation is only for z and is equivalent to

$$1 = \frac{z^2}{1+z} \iff z^2 - z - 1 = 0 \Rightarrow \boxed{z = \frac{1}{2}(1+\sqrt{5})}$$

We also observe that $y=0$ solves the second equation.

Then, with $y=0$, first equation implies that $\boxed{x = \frac{4}{3}}$

Problem 2.2.

a) Contollability matrix is $\mathcal{C} = [B \,\vdots\, AB] = \begin{bmatrix} 1 & 1 \\ 1 & a \end{bmatrix}$. So

controllable if and only if $a \neq 1$.

b) For $x = 0$, $v(x) = 0$. So from now on assume $x \neq 0$.

If $a \neq 1$, (A, B) is contollable, therefore for any initial data x,

there exists a contool $u_0 \in \mathbb{R}$ (depending on x) so that $x_1 = 0$.

Hence, if $a \neq 1$ we have $v(x) < \infty$, because we use first

u_0 and then $u_k \equiv 0$. This yields $x_k \equiv 0$ and $J < \infty$.

Now consider the case $a=1$. Let $x_k=(v_k, s_k)$. Then,

$$\left.\begin{aligned} v_{k+1} &= v_k + u_k \\ s_{k+1} &= s_k + u_k \end{aligned}\right\} \Rightarrow (v-s)_{k+1} = (v-s)_k \quad \forall k.$$

Hence, $(v-s)_k = (v-s)_0$ for every $k=0,1\ldots$

Suppose $\rho<1$. Take $u=(0,\ldots)$. Then, $v_k=v_0$, $s_k=s_0$ and

$$J(u) = \sum \rho^k (v_k^2+s_k^2) = \sum \rho^k (v_0^2+s_0^2) < \infty.$$

Suppose $\rho\geq 1$. For any control sequence $u:=(u_0,u_1,\ldots)$ if $J(u)<\infty$, then $\rho^k v_k^2$ goes to zero as k approaches to infinity. In particular, v_k is close to zero. But as $(v-s)_k=(v-s)_0$, if $v_0\neq s_0$, then s_k is close to $(v-s)_0$ and $\sum \rho^k s_k = +\infty$.

The only case left is $\rho\geq 1$, $a=1$ and $v_0=s_0$. In this case we take $u_k = \frac{1}{2} v_k$. This yieds $v_k=s_k=2^{-k} v_0$ and $J<\infty$. So we conclude that

$$\boxed{v(x)=+\infty \quad \text{for any } \rho\geq 1, a=1 \text{ and } x=(v_0,s_0) \text{ with } v_0\neq s_0}$$

Problem 2.3. Set $J(x,u) := \sum_{k=0}^{\infty} x_k^2 + u_k^2$.

a. Then,

$$v(x) = \inf_{u} \left\{ (x_0^2 + u_0^2) + \sum_{k=1}^{\infty} (x_k^2 + u_k^2) \right\}$$

Set $\tilde{u} := (u_1, u_2, \dots)$. As $x_0 = x$, we have

$$v(x) = x^2 + \inf_{(u_0, \tilde{u})} \left\{ u_0^2 + J(x_1, \tilde{u}) \right\}$$

$$= x^2 + \inf_{u_0} \left\{ u_0^2 + \underbrace{\inf_{\tilde{u}} J(x_1, \tilde{u})}_{= v(x_1)} \right\}$$

Since $x_1 = u_0 x + 1$, with $u = u_0$,

$$\boxed{v(x) = x^2 + \inf_{u} \left\{ u^2 + v(ux+1) \right\}}$$

b. Use the control $u = (0, 0, \dots)$. Then $x_0 = 0$, $x_k = 1$ for all $k \geq 1$. This implies that

$$J(0, u) = \sum_{k=1}^{\infty} 2^{-k} [1^2 + 0^2] < \infty .$$

c. For $x \neq 0$, let $u_0 = -1/x$. Then, $x_1 = 0$. Hence,

$$v(x) \leq x + (1/x)^2 + v(0) < \infty .$$

Problem 2.4.

a) It is clear that $J(0, x, u) = x$ and hence $v(0, x) = x$.

For $n > 0$,

$$v(n,x) = \max_{u} \left\{ -\frac{1}{2}u_0^2 + \mathbb{E}\left[-\frac{1}{2}\sum_{k=1}^{n} u_k^2 + x_n \right] \right\}$$

Note that

$$\mathbb{E}\left[-\frac{1}{2}\sum_{k=1}^{n} u_k^2 + x_n \right] = \mathbb{E}\Big[\underbrace{\mathbb{E}\left[-\frac{1}{2}\sum_{k=1}^{n} u_k^2 + x_n \,\middle|\, x_1 \right]}_{J(n-1,x_1,\tilde{u})} \Big]$$

Therefore,

$$v(n,x) = \max_{u_0} \Big\{ -\frac{1}{2}u_0^2 + \mathbb{E}\big[\underbrace{\max_{\tilde{u}} J(n-1,x_1,\tilde{u})}_{v(n-1,x_1)} \big]$$

So, with $u_0 = u$,

$$\boxed{v(n,x) = \max_{u} \left\{ -\frac{1}{2}u^2 + \mathbb{E}\left[v(n-1,x_1) \right] \right.}$$

b). We substitute $v(x) = x + n/2$ on the right-hand side of the above equation:

$$\max_{u} \left\{ -\frac{1}{2}u^2 + \mathbb{E}\left[x + u(z_0+1) + \frac{1}{2}(n-1) \right] \right\}$$

$$= \max_{u} \left\{ -\frac{1}{2}u^2 + x + u\left[\mathbb{E}(z_0)+1\right] + \frac{1}{2}n - \frac{1}{2} \right\}$$

$$= \max_{u} \left\{ -\frac{1}{2}u^2 + x + u + \frac{1}{2}n - \frac{1}{2} \right\}$$

$$= x + \frac{1}{2}n - \frac{1}{2} + \underbrace{\max\left\{ u - \frac{1}{2}u^2 \right\}}_{=1/2}$$

$$= x + \frac{1}{2}n = v(n,x).$$

Hence, $v(x) = x + n/2$ solves the dynamic programming.

Therefore, it is equal to the value function.

c). The maximizer in the above calculation ($\max_u \{u - \frac{1}{2}u^2\}$)

is $u^* = 1$. Hence, $u^* \equiv 1$ is optimal.

Problem 2.5

a). Set $x(t) = (y(t), y'(t))^T$. Then,

$$x'(t) = \begin{bmatrix} y'(t) \\ \sin y(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 + y(t) \end{bmatrix} u(t)$$

b) Since $\sin(x) \approx x$ for $x \approx 0$, we replace $\sin(y(t))$

with $y(t)$. The other nonlinear term is $y(t)u(t)$

we replace it with zero. Hence,

$$\boxed{x'(t) \cong \begin{bmatrix} y'(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t).}$$

c) For $h > 0$ but small, set $x_k := x(kh)$, $u_k = u(kh)$. Then,

$$\begin{aligned} x_{k+1} &\cong x_k + h x'(kh) \\ &= x_k + h \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_k \\ &= A x_k + B u_k \end{aligned}$$

where

$$\boxed{A = \begin{bmatrix} 1 & h \\ h & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ h \end{bmatrix}}$$

d) The controllability matrix is given by

$$\mathcal{C} = [B \,\vdots\, AB] = \begin{bmatrix} 0 & h^2 \\ 1 & h \end{bmatrix}$$

is full rank. Hence, (A, B) is controllable.

Problem 2.6

a) Set $x(t) = [y(t), y'(t), y''(t)]^T$. Then,

$$\boxed{x'(t) = \begin{bmatrix} y'(t) \\ y''(t) \\ y'''(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ (y'(t))^3 + u(t) \end{bmatrix}}$$

b) For small y, $y^3 \approx 0$. Hence, for $y(t) \approx 0$,

$$\boxed{x'(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t).}$$

c) With $0 < h$ small, set $x_k = x(kh)$, $u_k = u(kh)$. Then,

$$x_{k+1} \cong x_k + h x'(kh)$$

$$= x_k + h \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_k$$

Hence, $x_{k+1} = Ax_k + Bu_k$ where

$$\boxed{A = \begin{bmatrix} 1 & h & 0 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix}}$$

d) The controllability matrix $\mathcal{C}$ is given by

$$\mathcal{C} = [B \,\vdots\, AB \,\vdots\, A^2B] = \begin{bmatrix} 0 & 0 & h^3 \\ 0 & h^2 & 2h^2 \\ h & h & h \end{bmatrix}$$

has full rank. Hence, it is controllable.

Problem 3.1. First note that $B_p \cup B_u \cup B_n = \Omega$ and they are disjoint.

a) $\underbrace{\mathbb{P}(B_p|A)}_{=0.7 \text{ (given)}} + \underbrace{\mathbb{P}(B_n|A)}_{=0.1 \text{ (given)}} + \mathbb{P}(B_u|A) = 1$.

Hence, $\boxed{\mathbb{P}(B_u|A) =} 1 - 0.7 - 0.1 = \boxed{0.2}$. Similarly,

$\boxed{\mathbb{P}(B_n|A^c) =} 1 - \mathbb{P}(B_p|A^c) - \mathbb{P}(B_u|A^c) = 1 - 0.1 - 0.6 = \boxed{0.3}$

b) $\mathbb{P}(B_p) = \mathbb{P}(B_p|A)\mathbb{P}(A) + \mathbb{P}(B_p|A^c)\mathbb{P}(A^c)$

$= (0.7)(0.05) + (0.1)(0.95) = 0.13$

$$\boxed{\mathbb{P}(A|B_p) =} \frac{\mathbb{P}(B_p|A)\mathbb{P}(A)}{\mathbb{P}(B_p)} = \frac{(0.7)(0.05)}{0.13} = \boxed{\frac{7}{26}}$$