1. Show that for any Ornstein-Uhlenbeck stationary sequence $X_n, -\infty < n < \infty$, with parameter a and any $m > k$, the best linear estimate of X_m given $X_k, X_{k-1}, X_{k-2}, \ldots$, is given by

$$\tilde{X}_{m,k} = a^{m-k} X_k.$$

(Hint: it is an orthogonal projection, normal equations)

Answer. $\tilde{X}_{m,k} = a^{m-k} X_k$ is the best linear estimate if and only if $X_m - \tilde{X}_{m,k} = X_m - a^{m-k} X_k \perp X_l$ for all $l \le k$ which means that

$$\mathbf{E}\left[(X_m - a^{m-k} X_k)\right] X_l = 0 \text{ for all } l \le k$$

or

$$\mathbf{E} X_m X_l = a^{m-k} \mathbf{E} X_k X_l \text{ for all } l \le k.$$

But indeed, $\mathbf{E} X_m X_l = a^{m-l}$ and $a^{m-k} \mathbf{E} X_k X_l = a^{m-k} a^{k-l} = a^{m-l}$.

2. Assume $X_n, n \ge 0$, is a real valued space continuous homogeneous Markov chain with transition kernel $p(r, u)$. Assume the observation

$$Y_n = \gamma(X_n, W_n), n \ge 0,$$

where W_n is a sequence of i.i.d. independent of $X_0, X_1, \ldots$. Assume that for every $r \in \mathbf{R}$, $\gamma(r, W_n)$ is a continuous r.v. whose pdf. $l(r, v)$ is known.

a) Show that the pair (X_n, Y_n) is a space continuous homogeneous Markov chain with transition kernel

$$q^{r,u}(s, v) = p(r, u) l(u, v).$$

Write the recursive equation for UFDF. (Hint: look at HW 12.2)

Answer. Since W_{n+1} does not depend on X, and $Y_{[0,n]}$,

$$\begin{aligned}
&\mathbf{E}\left[1_{X_{n+1} \le a} 1_{\gamma(X_{n+1}, W_{n+1}) \le b} | X_{[0,n+1]}, Y_{[0,n]}\right] \\
&= \mathbf{E}\left[1_{u \le a} 1_{\gamma(u, W_{n+1}) \le b} | X_{[0,n+1]}, Y_{[0,n]}\right]_{u = X_{n+1}} \\
&= 1_{X_{n+1} \le a} (\mathbf{E} 1_{\gamma(u, W_{n+1}) \le b})|_{u = X_{n+1}} = 1_{X_{n+1} \le a} \int_{-\infty}^{b} l(u, v) dv|_{u = X_{n+1}} \\
&= 1_{X_{n+1} \le a} \int_{-\infty}^{b} l(X_{n+1}, v) dv.
\end{aligned}$$

Therefore for every $n \geq 0, a, b \in \mathbf{R}$,

$$
\begin{aligned}
&\mathbf{P}\left(X_{n+1} \leq a, Y_{n+1} \leq b | X_{[0,n]}, Y_{[0,n]}\right) \\
&= \mathbf{E}\left[1_{X_{n+1} \leq a} 1_{\gamma(X_{n+1}, W_{n+1}) \leq b} | X_{[0,n]}, Y_{[0,n]}\right] \\
&= \mathbf{E}\left\{\mathbf{E}\left[1_{X_{n+1} \leq a} 1_{\gamma(X_{n+1}, W_{n+1}) \leq b} | X_{[0,n+1]}, Y_{[0,n]}\right] | X_{[0,n]}, Y_{[0,n]}\right\} \\
&= \mathbf{E}\left[\int_{-\infty}^{b} l(X_{n+1}, v) dv 1_{X_{n+1} \leq a} | X_{[0,n]}, Y_{[0,n]}\right]
\end{aligned}
$$

Since $Y_i = \gamma(X_i, W_i)$, and W is independent of X,

$$
\begin{aligned}
&\mathbf{E}\left[\int_{-\infty}^{b} l(X_{n+1}, v) dv 1_{X_{n+1} \leq a} | X_{[0,n]}, Y_{[0,n]}\right] \\
&= \mathbf{E}\left\{\mathbf{E}\left[\int_{-\infty}^{b} l(X_{n+1}, v) dv 1_{X_{n+1} \leq a} | X_{[0,n]}, W_{[0,n]}\right] | X_{[0,n]}, Y_{[0,n]}\right\} \\
&= \mathbf{E}\left\{\mathbf{E}\left[\int_{-\infty}^{b} l(X_{n+1}, v) dv \ 1_{X_{n+1} \leq a} | X_{[0,n]}\right] | X_{[0,n]}, Y_{[0,n]}\right\} \\
&= \int_{-\infty}^{a} \int_{-\infty}^{b} l(u, v) dv \ p(X_n, u) du.
\end{aligned}
$$

Summarizing we got

$$
\begin{aligned}
&\mathbf{P}\left(X_{n+1} \leq a, Y_{n+1} \leq b | X_{[0,n]}, Y_{[0,n]}\right) \\
&= \int_{-\infty}^{a} \int_{-\infty}^{b} l(u, v) dv \ p(X_n, u) du.
\end{aligned}
$$

Taking the mixed derivative of both sides we obtain the statement.

b) (observation of a noisy reflection) Assume W_n are independent standard normal and $Y_n = |X_n + W_n|$.Find $l(u, v)$.

Answer. For any $v \geq 0$,

$$
\begin{aligned}
\mathbf{P}\left(|r + W_n| \leq v\right) &= \mathbf{P}\left(-v \leq r + W_n \leq v\right) = \mathbf{P}\left(-v - r \leq W_n \leq v - r\right) \\
&= \Phi\left(v - r\right) - \Phi(-v - r),
\end{aligned}
$$

where Φ is cumulative distribution function of a standard normal. Differentiating in v we obtain

$$
l(r, v) = \rho(v - r) + \rho(-v - r), v \geq 0,
$$

where ρ is the pdf of a standard normal.

3. The signal is defined by

$$X_{n+1} = 0.1\cos(2X_n)\,h + 0.14\sqrt{h}V_{n+1}, 0 \le n \le N,$$

where V_n are independent standard normal r.v.. Assume $X_0 \sim N(0, 0.1^2)$.

We observe

$$U_{n+1} = U_n + \arctan(X_n)h + 0.04\sqrt{h}W_{n+1}, 0 \le n \le N,$$

where W_n are independent standard normal independent of V_n. Assume $U_0 = 0$. We want to have filtering estimates of X_n and $H_n = \sum_{k=1}^{n} X_{k-1}$.

Take the interval $[-1, 1]$ as the spatial domain with the grid of step 0.01 (consider all density functions on this spatial grid and use it for the integrals on $\mathbf{R}$ in the recursive formulas). Take $h = 0.01, N = 200$.

a) Plot on the same graph $X_n, U_n, \hat{X}_n, 0 \le n \le 200$;

b) Plot on the same graph $H_n, \hat{H}_n, 0 \le n \le 200$;

c) (*maximum posterior probability estimate).* Guessing: since all the distributions are conditionally Gaussian, the following estimate defined by

$$\hat{X}_n^{\max} = \arg\max \gamma_n(r),$$

which is certain r_n maximizing the unnormalized density $\gamma_n(r)$, should coincide with $\hat{X}_n$. Check this guess by plotting on the same graph $X_n, \hat{X}_n, \hat{X}_n^{\max}, 0 \le n \le 200$. (Note you can get r_n by applying max command in Matlab to the vector $\gamma_n(r)$). The r.v. $\hat{X}_n^{\max}$ is called *maximum posterior probability estimate.*

d) Plot in 3D the FDF

$$\pi_n^a = \frac{\gamma_n(a)}{\int_{\mathbf{R}} \gamma_n(r)dr}, -1 \le a \le 1, 0 \le n \le 200.$$

Answer. The signal and observation are space continuous. we considered in class (see **section 8.1** of posted notes) the following model:

$$\begin{aligned} X_{n+1} &= b(X_n) + d(X_n)V_{n+1}, \\ Y_{n+1} &= c(X_n) + \sigma W_{n+1}, Y_0 = 0, \end{aligned}$$

The problem above can be put into this scheme by introducing $Y_n = U_n - U_{n-1}, n \ge 1, Y_0 = 0$ (also, if U_n is observed we can observe $Y_n = U_n - U_{n-1}$ as well).

The following formulas were derived:

$$\begin{aligned}
\mathbf{E}\left[H_n|U_{[0,n]}\right] &= \mathbf{E}\left[H_n|Y_{[0,n]}\right] = \frac{\int_{\mathbf{R}} \gamma_n(H_n,r)dr}{\int_{\mathbf{R}} \gamma_n(r)dr},\\
\mathbf{E}\left[h(X_n)|U_{[0,n]}\right] &= \mathbf{E}\left[h(X_n)|Y_{[0,n]}\right] = \frac{\int_{\mathbf{R}} h(a)\gamma_n(a)da}{\int_{\mathbf{R}} \gamma_n(r)dr}\\
&= \int_{\mathbf{R}} h(a)\frac{\gamma_n(a)}{\int_{\mathbf{R}} \gamma_n(r)dr}da = \int_{\mathbf{R}} h(a)\pi_n^a\, da,
\end{aligned}$$

where $\gamma_n(H_n,r)$ and $\gamma_n(r)$ satisfy the recursive equations:

$$\gamma_{n+1}(H_{n+1},a) = \int_{\mathbf{R}} p(r,a)\lambda^{-1}(Y_{n+1},r)[\gamma_n(H_n,r) + r\gamma_n(r)]dr,$$

and

$$\gamma_{n+1}(a) = \int_{\mathbf{R}} p(r,a)\lambda^{-1}(Y_{n+1},r)\gamma_n(r)dr.$$

Here $p(r,u)$ is the transition kernel of X_n and

$$\lambda^{-1}(Y_{n+1},r) = \exp\left\{\frac{c(r)Y_{n+1}}{\sigma^2} - \frac{1}{2}\frac{c(r)^2}{\sigma^2}\right\}.$$

Note that $\gamma_0(H_0,r) = 0$ and $\gamma_0(r) =$pdf of X_0.

The file math508final_3.m defines a function math508final_3(N,h,M,L) that plots four graphs that we are asked, N is the end of time interval, h is equation parameter, the grid of size $2L/M$ is used on the space interval $[-L.L]$. The use:

>>math508final_2(200,0.01,200,1)

In the graphs:

a) X_n is blue, U_n is green and $\hat{X}_n$ is red;

b) H_n is blue, $\hat{H}_n$ is read;

c) X_n is blue, $\hat{X}_n^{\max}$ is green and $\hat{X}_n$ is red;

COMMENT. *In this model it turned out that indeed $\hat{X}_n^{\max}$ and $\hat{X}_n$ coincide. It was so not only because of some conditionally Gaussian distributions but also $\arctan(x)$ in the signal is odd function and the signal starts approximately at zero. For example, it is not so if we replace* $\arctan(x)$ *by* $|x|$ *or by* $\cos(x)$.

4. (Kalman-Bucy Filter) Assume the signal satisfies the equation:

$$X_t = 1 + \int_0^t aX_s ds + bV_t, 0 \leq \ t \leq 10,$$

and the observation process is

$$Y_t = \int_0^t X_s ds + BW_t, 0 \leq t \leq 10.$$

Simulate on the time grid of $\Delta t = 0.01$, and plot on the same graph the observation Y_t, signal X_t, and its estimate $\hat{X}_t, 0 \leq t \leq 10$, for

a) $a = 0.1, b = 1, B = 0.3$; b) $a = 0.1, b = 1, B = 10$; c) $a = -1, b = 1, B = 0.3$; d) $a = -1, b = 2, B = 0.3$.

Answer. We found that the best estimate $\hat{X}_t$ is given by the formula

$$\hat{X}_t = \alpha_t \left(x_0 + \int_0^t B^{-2} \alpha_s^{-1} P_s dY_s \right), t \geq 0,$$

where

$$\alpha_t = \exp \left\{ (a - B^{-2}\lambda_1)t \right\} \frac{(K+1)e^{t/C}}{Ke^{t/C}+1}, t \geq 0,$$

and $P_t = \mathbf{E}\left[\left(X_t - \hat{X}_t \right)^2 \right]$ is given by the formula

$$P_t = \lambda_1 + \frac{\lambda_2 - \lambda_1}{Ke^{t/C}+1}, t \geq 0,$$

where

$$\begin{aligned}
\lambda_1 &= \sqrt{a^2B^4 + b^2B^2} + aB^2, \\
\lambda_2 &= -\sqrt{a^2B^4 + b^2B^2} + aB^2, \\
C &= \frac{B^2}{\lambda_1 - \lambda_2}, \\
K &= \frac{-\lambda_2}{\lambda_1}.
\end{aligned}$$

Note that the mean square error P_t is increasing function of t. Obviously, $P_0 = 0$ and $\lim_{t\to\infty} P_t = \sup_{t\geq 0} P_t = \lambda_1 = \sqrt{a^2B^4 + b^2B^2} + aB^2$:

in any case $P_t \leq \lambda_1 \leq bB$ for all t.

The file math508final_4.m defines a function math508final_4(n, T, a, b, B, x_0) that can plot the graphs that we are asked, T is the end of time interval, $T = n \cdot (\Delta t), a, b, B$ are the coefficients in the equation, x_0 is the starting deterministic point. The use:

```
>>math508final_4(1000,10,0.1,1,0.3,1)
```

etc.

In the graphs: X_n is blue, Y_n is green and $\hat{X}_n$ is red