Stochastic analysis, autumn 2011, Exercises-3, Solutions 27.09.11

1.

- Show that the linear space generated by the Haar system, which coincides with the set of functions which are piecewise constant on the dyadic partition D_n for some $n \in \mathbb{N}$, is dense in the space of continuous functions $C([0,1],\mathbb{R})$ under the supremum norm

$$\parallel f \parallel_\infty := \sup_{t\in[0,1]} |f(t)|$$

A continuous function f is uniformly continuous on a compact, therefore for every $\varepsilon > 0$ $\exists n(\varepsilon)$ such that

$$|f(t) - f(k2^{-n})| < \varepsilon \qquad \text{when } t \in [k2^{-n}, (k+1)2^{-n})$$

and

$$\int_0^1 (f(t) - f_{n(\varepsilon)}(t))^2 dt \le \varepsilon^2$$

We recall Luzin's theorem from real analysis: if $x : [0,1] \to \mathbb{R}$ is a measurable function, for all $\varepsilon > 0$ there exists a continuous function f such that

$$\lambda(\{t : x(t) \neq f(t)\}) < \varepsilon$$

where $\lambda(dt)$ is Lebesgue measure.

- Show that $C([0,1],\mathbb{R})$ is dense in $L^2([0,1], dt)$.

Let $X(\omega) \in L^2(\Omega, \mathcal{F}, P)$. Let

$$X_n(\omega) = X(\omega)\mathbf{1}(|X(\omega)| \le n)$$

Then

$$\lim_{n\uparrow\infty} X_n(\omega)^2 \uparrow X(\omega)^2 \quad \forall\omega$$

as $n \uparrow \infty$ by the dominated convergence theorem

$$\lim_{n\uparrow\infty} \int_0^1 (X(\omega) - X_n(\omega))^2 P(d\omega) = 0$$

Bounded random variables are in dense in $L^2(\Omega, \mathcal{F}, P)$. In particular bounded measurable functions are dense in $L^2([0,1], dt)$. It is enough to show that a bounded measurable function can be approximated by a continuous function in $L^2([0,1], dt)$.

Let

$$\sup_{t\in[0,1]} |x(t)| = \parallel x \parallel_\infty < \infty$$

By Luzin theorem for each n there is a bounded continuous function $y_n(t)$ such that $|y_n(t)| \le \parallel x \parallel_\infty$ and

$$\lambda(\{t : x(t) \neq y_n(t)\}) < 1/n$$

Then

$$\int (x(t) - y_n(t))^2 dt \le 2 \parallel x \parallel_\infty \lambda(\{t : x(t) \neq y_n(t)\}) \to 0$$

- Show that the Haar system is a complete orthonormal basis of $L^2([0,1],dt)$.
  Consider $\dot\eta_d(t)$ $\dot\eta_{d'}(t)$ the Haar functions defined in the lectures with $d \in D_n \setminus D_{n-1}$, $d' \in D_{n'} \setminus D_{n'-1}$. There are three possibilities:
  $d' = d$ and we see that $\int \dot\eta_d(t)^2 dt = 2^n 2^{-n} = 1$.
  Otherwise either the $n = n'$ and $d \neq d'$, in such case $\dot\eta_d$ and $\dot\eta'_d$ have disjoint support,
  or $n \neq n'$, for example $n < n'$, and also in such case either $\dot\eta'_d$ have disjoint support, or the $\dot\eta_d$ is constant on the support of $\dot\eta_{d'}$, and

  $$\int \dot\eta_{d'}(t)dt = 0$$

  since $\dot\eta_{d'}$ is antisymmetric around d'.
  Now every function $f \in C([0,1])$ is uniformly continuous on the compact $[0,1]$. Therefore we can always approximate in supremum norm f by a sequence of functions f_n which are piecewise constant on intervals with extremes in D_n, and such function is expressed as linear comnbinations of $\dot\eta_d$ functions with $d \in D_n$
  To conclude the argument, just note that the supremum norm is stronger than the norm $L^2([0,1],dt)$, so that the sequence f_n appoximates f also in L^2-sense.

2. Let $X(\omega) \in \mathbb{R}$ and $Y(\omega) = (Y_1(\omega),\dots,Y_d(\omega)) \in \mathbb{R}^d$ with $X, Y_i \in L^2(\Omega,\mathcal{F},P)$.
   Consider the linear subspace generated by $Y(\omega)$

   $$\mathrm{Lin}(Y) = \{a + b\cdot Y(\omega) : a \in \mathbb{R}, b \in \mathbb{R}^d\}.$$

   Note that this is a $(d+1)$-dimensional space.
   We define the best linear estimator of X given Y as the L^2-orthogonal projection $\hat{E}(X|Y)$ of X on the linear subspace $\mathrm{Lin}(Y)$ generated by Y.
   Equivalently

   $$\hat{E}(X|Y)(\omega) = \hat{a} + \hat{b}\cdot Y(\omega)$$

   for some deterministic $\hat{a} \in \mathbb{R}$ $\hat{b} \in \mathbb{R}^d$ where

   $$(\hat{a}, \hat{b}Y(\omega)) = \arg\min_{a,b} E\big(\{X - (a + b\cdot Y)\}^2\big)$$

   Note that the conditional expectation $E(X|Y) = E(X|\sigma(Y))$ is the L^2-orthogonal projection of X on the infinite dimensional subspace $L^2(\Omega,\sigma(Y),P) \supset \mathrm{Lin}(Y)$, and in general $E(X|Y) \neq \hat{E}(X|Y)$.

   - Show that

     $$\hat{E}(X|Y) = E(X) + (Y - E(Y))\Sigma_{YY}^{-1}\Sigma'_{XY}$$
     $$E\Big((X - \hat{E}(X|Y))^2\Big) = \Sigma_{XX}\Sigma_{YY}^{-1}\Sigma'_{XY}$$

     where the covariance matrix of $(X,Y) = (X, Y_1,\dots,Y_d)$ is denoted as

     $$\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma'_{XY} & \Sigma_{YY} \end{pmatrix}$$

Hint: assume $E(X) = E(Y_i) = 0$, and maximize the mean square error with respect to the parameters a, b.

- Show that when the vector (X, Y) is jointly gaussian, all conditional distributions are gaussian and the best linear estimator $\hat{E}(X|Y)$ coincides with the conditional expectation $E(X|Y)$. (Use Bayes formula!).

  Hint: recall that the joint distribution of a gaussian vector is specified by the mean vector and covariance matrix.

R. Assume that $E(Y) = 0$ since Y and $Y - E(Y)$ generate the same linear subspace, and $E(X) = 0$, otherwise project first $(X - E(X))$ where $E(X) \in \text{Lin}(Y)$. In dimension 1

$$E\big(\big\{X - (a + b \cdot Y)\big\}^2\big) = E(X^2) + a^2 + bE(Y^\top Y)b^\top + 2ab \cdot E(Y) - 2aE(X) - 2b \cdot E(XY) =$$
$$E(X^2) + a^2 + bE(Y^\top Y)b^\top - 2b \cdot E(XY)$$

We obtain

$$0 = \frac{\partial}{\partial a} E\big(\big\{X - (a + b \cdot Y)\big\}^2\big) = 2a,$$
$$0\frac{\partial}{\partial b_i} E\big(\big\{X - (a + b \cdot Y)\big\}^2\big) = 2(bE(Y^\top Y) - E(XY)) \Longleftrightarrow a = 0, \quad b = \; E(XY)E(Y^\top Y)^{-1}$$

We obtain the best linear estimator

$$\hat{E}(X|Y) = E(X) + E(XY)E(Y^\top Y)^{-1}(Y - E(Y))$$

Still assuming $E(X) = 0$ $E(Y) = 0$, by orthogonality

$$E\Big(\hat{E}(X|Y)(X - \hat{E}(X|Y))\Big) = 0$$

and

$$\text{Var}(X) = E(X^2) = E\Bigg(\Bigg\{\hat{E}(X|Y) + (X - \hat{E}(X|Y))\Bigg\}^2\Bigg) =$$
$$E(\hat{E}(X|Y)^2) + E((X - \hat{E}(X|Y))^2)$$

that is

$$E(\hat{E}(X|Y)^2) = E(X^2) - E(\hat{E}(X|Y)^2) = E\Bigg(\Bigg\{X - \hat{E}(X|Y)^2\Bigg\}^2\Bigg) =$$
$$E(X^2) - E(XY)E(Y^\top Y)^{-1}E(YY^\top)E(YY^\top)^{-1}E(XY^\top) = E(X^2) - E(XY)E(Y^\top Y)^{-1}E(XY^\top)$$

Next we compute the conditional distributions of a multivariate gaussian.

Let $(X, Y) \sim \mathcal{N}(0, \Sigma)$. We assume that $E(X) = 0$ and $E(Y) = 0$Âăotherwise we shift the gaussian distribution considering the pairs $X' = (X - E(X))$ and $Y' = (Y - E(Y))$.

Denote the **precision matrix** $D = \Sigma^{-1}$, $X(\omega) \in \mathbb{R}^{n_x}, Y(\omega) \in \mathbb{R}^{n_y}$, $n = n_x + n_y$ where Σ is the covariance matrix of (X, Y).

By Bayes' formula

$$p_{XY}(x,y) = (2\pi)^{-n/2}\sqrt{|D|}\exp\left(-\frac{1}{2}\left\{(x,y)D(x,y)^\top\right\}\right) = p_X(x)p_{Y|X}(y|x)$$

where

$$p_X(x) = (2\pi)^{-n_x/2}|\Sigma_{xx}|^{-1/2}\exp\left(-\frac{1}{2}\left\{x\Sigma_{xx}^{-1}x^\top\right\}\right)$$

$$p_{Y|X}(y|x) = \frac{p_{XY}(x,y)}{p_X(x)} = (2\pi)^{-n_y/2}\sqrt{|D|\times|\Sigma_{xx}|}\exp\left(-\frac{1}{2}\left\{(x,y)D(x,y)^\top - x\Sigma_{xx}^{-1}x^\top\right\}\right) =$$
$$(2\pi)^{-n_y/2}\sqrt{|D|\times|\Sigma_{xx}|}\exp\left(-\frac{1}{2}\left\{x(D_{xx}-\Sigma_{xx}^{-1})x^\top\right)\exp\left(-\frac{1}{2}\left\{yD_{yy}y^\top + 2yD_{yx}^\top x^\top\right\}\right) =$$
$$(2\pi)^{-n_y/2}\sqrt{|D|\times|\Sigma_{xx}|}\exp\left(-\frac{1}{2}\left\{x(D_{xx}-D_{xy}D_{yy}^{-1}D_{xy}^\top - \Sigma_{xx}^{-1})x^\top\right\}\right)$$
$$\times\exp\left(-\frac{1}{2}\left\{(y+xD_{xy}D_{yy}^{-1})D_{yy}(y+xD_{xy}D_{yy}^{-1})^\top\right\}\right)$$

Now conditionally on X we treat x as a constant it follows that the conditional distribution $p_{Y|X}(y|x)$ is gaussian with conditional covariance matrix

$$\Sigma_{y|x} = D_{yy}^{-1} \tag{1}$$

and conditional mean

$$E(y|x) = -xD_{xy}D_{yy}^{-1} \tag{2}$$

Also since this conditional variance does not depend on x we must have

$$\Sigma_{xx}^{-1} = D_{xx} - D_{xy}D_{yy}^{-1}D_{xy}^\top \tag{3}$$

and also

$$|D_{yy}| = |D|\times|\Sigma_{xx}| = |\Sigma_{xx}|/|\Sigma| \tag{4}$$

Note also that by inverting the roles of Σ and D ($D = \Sigma^{-1}$ is also a symmetric non-negative matrix, which corresponds to a covariance matrix), we obtain

$$D_{xx}^{-1} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}^\top, \tag{5}$$
$$|\Sigma_{yy}| = |\Sigma|\times|D_{xx}| = |D_{xx}|/|D| \tag{6}$$

By changing the roles of x and y we obtain also

$$\Sigma_{x|y} = D_{xx}^{-1}, \tag{7}$$
$$\Sigma_{yy}^{-1} = D_{yy} - D_{xy}^\top D_{xx}^{-1}D_{xy}, \tag{8}$$
$$D_{yy}^{-1} = \Sigma_{yy} - \Sigma_{xy}^\top\Sigma_{xx}^{-1}\Sigma_{xy} = \Sigma_{y|x} \tag{9}$$

Now we use the property of the inverse matrix: since $\Sigma D = D\Sigma = Id$

$$\begin{pmatrix}\Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^\top & \Sigma_{yy}\end{pmatrix}\begin{pmatrix}D_{xx} & D_{xy} \\ D_{xy}^\top & D_{yy}\end{pmatrix} = \begin{pmatrix}Id & 0 \\ 0 & Id\end{pmatrix} = \begin{pmatrix}D_{xx} & D_{xy} \\ D_{xy}^\top & D_{yy}\end{pmatrix}\begin{pmatrix}\Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^\top & \Sigma_{yy}\end{pmatrix} \quad (10)$$

we have

$$\Sigma_{xx}D_{xx} + \Sigma_{xy}D_{xy}^\top = Id = D_{xx}\Sigma_{xx} + D_{xy}\Sigma_{xy}^\top \quad (11)$$
$$\Sigma_{xx}D_{xy} + \Sigma_{xy}D_{yy} = 0 = D_{xx}\Sigma_{xy} + D_{xy}\Sigma_{yy} \quad (12)$$
$$\Sigma_{xy}^\top D_{xx} + \Sigma_{yy}D_{xy}^\top 0 = D_{xy}^\top \Sigma_{xx} + D_{yy}\Sigma_{xy}^\top \quad (13)$$
$$\Sigma_{xy}^\top D_{xy} + \Sigma_{yy}D_{yy}^\top = Id = D_{xy}^\top \Sigma_{xy} + D_{yy}\Sigma_{yy}^\top \quad (14)$$

we can use it to obtain a new expression for the conditional expectation (2) : by using (9) ,(11),(12)

$$\begin{aligned}E(y|x) &= -xD_{xy}D_{yy}^{-1} = -xD_{xy}\big(\Sigma_{yy} - \Sigma_{xy}^\top\Sigma_{xx}^{-1}\Sigma_{xy}\big) \\ &= x\big(-D_{xy}\Sigma_{yy} + D_{xy}\Sigma_{xy}^\top\Sigma_{xx}^{-1}\Sigma_{xy}\big) \\ &= x\big(D_{xx}\Sigma_{xy} + \{Id - D_{xx}\Sigma_{xx}\}\Sigma_{xx}^{-1}\Sigma_{xy}\big) \\ &= x\big(D_{xx}\Sigma_{xy} + \Sigma_{xx}^{-1}\Sigma_{xy} - D_{xx}\Sigma_{xy}\big) = x\Sigma_{xx}^{-1}\Sigma_{xy}\end{aligned}$$

By changing the roles of x and y we get also

$$E(x|y) = -yD_{xy}^\top D_{xx}^{-1} = y\Sigma_{yy}^{-1}\Sigma_{xy}^\top \quad (15)$$

When X and Y a priori have non zero mean, by using $X' = (X - E(X))$ and $Y' = (Y - E(Y))$ we obtain

$$E(X|Y) = E(X) + \{Y - E(Y)\}\Sigma_{yy}^{-1}\Sigma_{xy}^\top \quad (16)$$
$$E(Y|X) = E(Y) + \{X - E(X)\}\Sigma_{xx}^{-1}\Sigma_{xy} \quad (17)$$

It follows also that

$$D_{xy} = -\Sigma_{xx}^{-1}\Sigma_{xy}D_{yy} = -\Sigma_{xx}^{-1}\Sigma_{xy}\Sigma_{y|x}^{-1} = -\Sigma_{x|y}^{-1}\Sigma_{xy}\Sigma_{yy}^{-1} \quad (18)$$

and

$$D = \begin{pmatrix}D_{xx} & D_{xy} \\ D_{xy}^\top & D_{yy}\end{pmatrix} = \Sigma^{-1} = \begin{pmatrix}\Sigma_{x|y}^{-1} & -\Sigma_{x|y}^{-1}\Sigma_{xy}\Sigma_{yy}^{-1} \\ -\Sigma_{xx}^{-1}\Sigma_{xy}\Sigma_{y|x}^{-1} & \Sigma_{y|x}^{-1}\end{pmatrix} \quad (19)$$

3. Let $(B_t(\omega) : t \in [0,1])$ a Brownian motion, and $D_n = (k2^{-n} : k = 0, 1, \dots, 2^n)$.

   Show that for fixed n and dyadic indexes

   $d = (2k+1)2^{-n} \in D_n \setminus D_{n-1}, d_- = 2k2^{-n}, d_+ = (2k+2)2^{-n} \in D_{n-1}$

   with $k = 0, \dots 2^{n-1}$,

$$G_d(\omega) := \left(B_d(\omega) - \frac{B_{d-}(\omega) + B_{d+}(\omega)}{2}\right)2^{(n+1)/2}, \quad d \in D$$

are i.i.d. standard gaussian variable ($E(G_d)=0, E(G_d^2)=0$).

R.

$$G_d(\omega)=2^{(n-1)/2}\Big(B_d(\omega)-B_{d-}(\omega)\Big)-2^{(n-1)/2}\Big(B_{d+}(\omega)-B_d(\omega)\Big)$$

where the intervals $(d_-,d)\cap(d,d_+)=\emptyset$ and also the intervals $(d_-,d_+)\cap(d'_-,d'_+)=\emptyset$ when $d'\neq d\in D_n\setminus D_{n-1}$. Therefore the increments on the right hand side are independent, and also are independent from G'_d where $d'\neq d\in D_n\setminus D_{n-1}$

Since $d-d_-=d_+-d=2^{-n}$, it follows that G_d is gaussian with $E(G_d)=0$ and variance $E(G_d^2)=2\big(2^{(n-1)/2}\big)^2 2^{-n}=1$.

4. Let $G(\omega)\sim\mathcal{N}(0,1)$, and $f\in L^2(\mathbb{R},d\gamma)$ where $\gamma(dx)=\phi(x)dx$.

   Here

$$\phi(x)=\frac{1}{\sqrt{2\pi}}\exp\big(-\frac{x^2}{2}\big)$$

   denotes the standard gaussian density.

   Consider the function

$$u(t,x)=E_P\Big(f(x+G\sqrt{t})\Big)$$

   - Show that $u(t,x)$ is smooth in the open set $(0,\infty)\times\mathbb{R}$. This does not require any smoothness on f.

     Hint: write

$$\frac{u(t+\varepsilon,x)-u(t,x)}{\varepsilon},\quad \frac{u(t,x+\varepsilon)-u(t,x)}{\varepsilon}$$

     as integrals, and do an opportune change of variable in order to use the smoothness of the gaussian density ϕ when you take the limit as $\varepsilon\to 0$.

$$\begin{aligned}\frac{u(t,x+\varepsilon)-u(t,x)}{\varepsilon}&=\frac{1}{\varepsilon}E\Big(f(x+\varepsilon+G\sqrt{t})-f(x+G\sqrt{t})\Big)=\\ &=\frac{1}{\varepsilon}\int_{\mathbb{R}}\Big(f(x+\varepsilon+y\sqrt{t})-f(x+y)\Big)\phi(y)dy\\ &=\frac{1}{\varepsilon}\int_{\mathbb{R}}f(x+y\sqrt{t})\Big(\phi\Big(y-\frac{\varepsilon}{\sqrt{t}}\Big)-\phi(y)\Big)dy\\ &\longrightarrow\int_{\mathbb{R}}f(x+y\sqrt{t})\frac{(-1)}{\sqrt{t}}\phi'(y)dy=\int_{\mathbb{R}}f(x+y\sqrt{t})\frac{1}{\sqrt{t}}\phi(y)dy\\ &=\frac{1}{\sqrt{t}}E\big(f(x+G\sqrt{t})G\big)=\frac{1}{t}E\big(f(x+G\sqrt{t})G\sqrt{t}\big)\end{aligned}$$

     For the time-derivative,

$$
\begin{aligned}
&\frac{u(t+\varepsilon,x)-u(t,x)}{\varepsilon}=\frac{1}{\varepsilon}E\Big(f(x+G\sqrt{t+\varepsilon})-f(x+G\sqrt{t})\Big)=\\
&=\frac{1}{\varepsilon}\int_{\mathbb{R}}f(x+y)\Big(\frac{1}{\sqrt{t+\varepsilon}}\phi\Big(\frac{y}{\sqrt{t+\varepsilon}}\Big)-\frac{1}{\sqrt{t}}\phi\Big(\frac{y}{\sqrt{t}}\Big)\Big)dy\\
&\longrightarrow\int_{\mathbb{R}}f(x+y)\frac{\partial}{\partial t}\Big(\frac{1}{\sqrt{t}}\phi\Big(\frac{y}{\sqrt{t}}\Big)\Big)dy=\int_{\mathbb{R}}f(x+y)\frac{1}{2t}\frac{1}{\sqrt{t}}\phi\Big(\frac{y}{\sqrt{t}}\Big)\Big(\frac{y^2}{t}-1\Big)dy\\
&=\frac{1}{2t}E\Big(f(x+G\sqrt{t})(G^2-1)\Big)=\frac{1}{2t^2}E\Big(f(x+G\sqrt{t})(G^2t-t)\Big)
\end{aligned}
$$

where we have computed the partial derivative

$$
\begin{aligned}
&\frac{\partial}{\partial t}\Big(\frac{1}{\sqrt{t}}\phi\Big(\frac{y}{\sqrt{t}}\Big)\Big)=-\frac{1}{2}t^{-3/2}\phi\Big(\frac{y}{\sqrt{t}}\Big)+t^{-1/2}\phi'\Big(\frac{y}{\sqrt{t}}\Big)\Big(-\frac{y}{2}t^{-3/2}\Big)\Big)=\\
&\frac{(-1)}{2}t^{-3/2}\Big(\phi\Big(\frac{y}{\sqrt{t}}\Big)+yt^{-1/2}\phi'\Big(\frac{y}{\sqrt{t}}\Big)\Big)=\\
&\frac{-1}{2}t^{-3/2}\Big(\phi\Big(\frac{y}{\sqrt{t}}\Big)-yt^{-1/2}\phi\Big(\frac{y}{\sqrt{t}}\Big)\frac{y}{\sqrt{t}}\Big)=\frac{1}{2}t^{-3/2}\phi\Big(\frac{y}{\sqrt{t}}\Big)\Big(\frac{y^2}{t}-1\Big)
\end{aligned}
$$

Taking the limit as $\varepsilon\to 0$ inside the expectation needs to be justified. It follows by an uniformly integrability condition which will be discussed this later in the course. It holds when $x\mapsto f(x)$ has polynomial growth, since the gaussian distribution has all exponential moments $E(\exp(\theta G))<\infty$ for $\theta\in\mathbb{R}$.

- Use the gaussian integration by parts formula to express the partial derivatives for $t>0$

$$
\frac{\partial}{\partial t}u(t,x),\quad \frac{\partial}{\partial x}u(t,x),\quad \frac{\partial^2}{\partial x^2}u(t,x)
$$

R We have since $B_t\sim G\sqrt{t}$ where $\sim$ means identity in law and G is a standard gaussian,

$$
\begin{aligned}
\frac{\partial}{\partial t}u(t,x)&=\frac{1}{2t^2}E\Big(f(x+B_t)(B_t^2-t)\Big),\\
\frac{\partial}{\partial x}u(t,x)&=\frac{1}{t}E\big(f(x+B_t)B_t\big)
\end{aligned}
$$

For the second derivative w.r.t. x we apply the integration by parts formula to the first derivative

$$
\begin{aligned}
&\frac{\partial^2}{\partial x^2}E(f(x+B_t))=\frac{1}{t}\frac{\partial}{\partial x}E(f(x+B_t)B_t)\\
&=\frac{1}{t}\Big\{\frac{\partial}{\partial x}E(f(x+B_t)(x+B_t))-\frac{\partial}{\partial x}\Big(xE\big(f(x+B_t)\big)\Big)\Big\}
\end{aligned}
$$

for the first term we apply the gaussian integration by parts to $x\mapsto g(x):=$

$f(x)x$,

$$\frac{\partial^2}{\partial x^2}E(f(x+B_t)) = \frac{1}{t}\left\{\frac{1}{t}E_P(g(x+B_t)B_t) - E(f(x+B_t)) - \frac{x}{t}E(f(x+B_t)B_t)\right\} =$$
$$\frac{1}{t}\left\{\frac{1}{t}E_P(f(x+B_t)(x+B_t)B_t) - E(f(x+B_t)) - \frac{x}{t}E(f(x+B_t)B_t)\right\}$$
$$= \frac{1}{t^2}E\big(f(x+B_t)(B_t^2-t)\big) = 2\frac{\partial}{\partial t}E(f(x+B_t))$$

We see that $u(x,t) = E(f(x+B_t))$ satisfies the partial differential equation

$$\frac{\partial}{\partial t}u(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,x)$$

5. Let

$$p(x,t) = \frac{1}{\sqrt{t}}\phi\left(\frac{x}{\sqrt{t}}\right)$$

- By using the Markov property of Brownian motion (which follows from the independence of increments), show that for $0 \le t \le T$

  $$p(y-x, T-t)dy = P(B_T \in dy | B_t = x) = P(B_{T-t} + x \in dy)$$

  R. For a non-negative measurable test function $f(x,y)$

  $$E(f(B_T, B_t)) = E(f(B_t + (B_T - B_T), B_t)) = E(f(B_t + G\sqrt{T-t}, B_t))$$
  $$= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(x+y,x)\frac{1}{\sqrt{T-t}}\phi\left(\frac{y}{\sqrt{T-t}}\right)dy\right)\frac{1}{\sqrt{t}}\phi\left(\frac{x}{\sqrt{t}}\right)dx$$

  Denote for $t \in [0,T]$

  $$v(t,x) = \int_{\mathbb{R}} f(y)p(y-x, T-t)dy = E_P(f(B_T)|B_t = x)$$

  for some $f \in L^2(\mathbb{R}, d\gamma)$, where $\gamma(dy) = p(y,T)dy$ is the $\mathcal{N}(0,T)$ gaussian measure.

- Show that $v(t,x)$ is smooth in $[0,T)\times\mathbb{R}$ with respect to the variables (t,x), the partial derivatives

  $$\frac{\partial}{\partial t}v(t,x), \quad \frac{\partial}{\partial x}v(t,x), \quad \frac{\partial^2}{\partial x^2}v(t,x)$$

  R. We have

  $$v(t,x) = E_P(f(B_T)|B_t = x) = E_P(f(x + B_T - B_t)|B_t = x) = E_P(f(x + B_T - B_t)) =$$
  $$E_P(f(x + G\sqrt{T-t})) = u(T-t, x)$$

  Therefore

  $$\frac{\partial}{\partial t}v(t,x) = \frac{\partial}{\partial t}u(T-t,x) = -\frac{\partial u}{\partial t}(T-t,t)$$
  $$\frac{\partial}{\partial x}v(t,x) = \frac{\partial}{\partial x}v(T-t,x) = \frac{\partial v}{\partial x}(T-t,x)$$
  $$\frac{\partial^2}{\partial x^2}v(t,x) = \frac{\partial^2}{\partial x^2}v(t,x) = \frac{\partial^2 v}{\partial x^2}(t,x)$$

$v(t,x)$ satifies the partial differential equation (heath equation)

$$\frac{\partial}{\partial t}v(t,x) = -\frac{\partial u}{\partial t}(T-t,t) = -\frac{1}{2}\frac{\partial^2 u}{\partial x^2}(T-t,t) = -\frac{\partial^2}{\partial x^2}u(T-t,t) \quad ,0 \leq t < T$$