COMP 767 Winter 2018
Raihan Seraj
Student ID: 260752605

---

1. Question 1:

   (a) Explain what is an oblique projection

   Ans: An oblique projection is an example of non orthogonal projection. In case of orthogonal projection some approximation of $\hat{v}$ is obtained by projecting v on to the span of Φ, where Φ is the feature matrix so as to minimize $\hat{v} \to \|\hat{v} - v\|$. Thus, this approximation is obtained by projecting v orthogonally with respect to the $\|.\|$. In case of an oblique projection, it is any projection Π onto span ϕ, that is any linear operator that satisfies the following $\Pi^2 = \Pi$ and whose range is $span(\phi)$. In general a simple example of non-orthogonal(oblique projection is)

$$P^2 = \begin{bmatrix} 0 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \alpha & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \alpha & 1 \end{bmatrix} = P$$

   where the projection is orthogonal if and only if $\alpha = 0$

   (b) Show projected bellman operator gives rise to an oblique projection

   Ans: For an oblique projection, the following can be written $\Pi_x = \phi \pi_x$ where $\pi_x = (X'\Phi)^{-1}X'$. Now considering the Bellman operator τ and Recalling that $L = I - \gamma P$ and $v = \tau v$ is equal to $L^{-1}r$ is the unique fixed point of the bellman operator. Thus we can write the following:

$$\hat{v}_x = \Phi w_X$$

   The above expression holds because we are assuming $\{\Phi_1 ... \Phi..m\}$ forms linear independent sets. Thus the fixed point equation is

$$\Phi w_x = \Pi_X(r + \gamma \Phi w_x$$

   Multiplying both sides by π_x we have the following

$$\begin{aligned} w_X &= \pi_X(I - \gamma \pi_X P \phi)^{-1} \pi_X r \\ where \quad \pi_X &= (X'\Phi)^{-1}X' \end{aligned}$$

Now while substituting the expression we obtain the following

$$\begin{aligned} w_X &= (I - \gamma(X'\Phi)^{-1}X'P\Phi)^{-1}(X'\Phi)^{-1}X'r \\ &= [(X'\Phi)(I - \gamma(X'\Phi)^{-1}X'P\Phi)]^{-1}X'r \\ &= (X'(I - \gamma P)\Phi)^{-1}X'r \\ &= (X'L\Phi)^{-1}X'Lv \\ &= \pi_{L'X}v \end{aligned}$$

where we substituted the fact that $r = Lv$.

The principle of the Bellman Residual method is to look for $\hat{v} \in span(\Phi)$ so that it minimizes norm of the Bellman Residual, since $\hat{v}$ is of the form Φw it can be seen that Error at the bellman residual $E_{BR}(\hat{v} = \|\Phi w - \gamma P \Phi w - r\|_\zeta$
$= \|\Psi w - r\|_\zeta$ where $\hat{v}_{BR} = \Phi w_{BR}$ with $w_{BR} = (\Psi'\Xi\Psi)^{-1}\Psi'\Xi r$

From the above expressions we can see that the solution of the fixed point of the bellman residual exists when also the projected bellman operator is an oblique projection. This means that when the projection is an oblique projection, there exist a solution of the projected equation where the projection of v onto $span(\Phi)$ orthogonally to span $L'X$ that is formally $\hat{v} = \Pi_{L'X}v$

(c) How does oblique projection change with λ Ans:In an oblique projection, θ is projected in such a way so that it is perpendicular to the new subspace $\mathcal{L}$ and an oblique projection to its original subspace K thus due to oblique projection $b - A\hat{\theta} = 0$ therefore when λ changes the planes rotate that is this will likely cause a change in the angle between the original subspace $\mathcal{K}$ and the new subspace $\mathcal{L}$

2. Question 2

   (a) Explain what is Galerkin's method

   Ans: Galerkin methods are a class of methods for converting a continuous operator problem to a discrete problem. In practice it is equivalent to applying the method of variation of parameter to a function space by converting equation to a weak formulation.

   (b) Explain why the projected Bellman Operator can be interpreted as Galerkin's method?

   Ans: The general idea of Galerkin method is to approximate the fixed point $\tau V^* = V^*$ where τ is the Bellman Operator. The projected method of solving the this equation consists of two finite dimensional linear subspaces $\mathcal{F}$ and $\mathcal{G}$ of a Banach space $B\mathcal{B}$ of function over $\mathcal{X}, \mathcal{F}, \mathcal{G}$ sharing a common space $d \in \mathcal{N}$ and then solving the projected equation

$$\Pi_{\mathcal{G}} TV = \Pi_{\mathcal{G}} V \quad for \quad V \in \mathcal{F}$$

   where $\Pi_{\mathcal{G}} : \mathcal{B} \Rightarrow \mathcal{F}$ is the projection operator. Since this leads to a $d \times d$ linear system of equations, hence once the bases of $\mathcal{F}$ and $\mathcal{G}$ is fixed $\Pi_{\mathcal{G}}$ is the corresponding orthogonal projection and $T : \mathcal{F} \to \mathcal{B}$ is bounded and the projected bellman operator can be interpreted as a Galerkin's method where

$$< TV, g >=< V, g > \quad for \quad all \quad g \in \mathcal{G}$$

3. Question 3

   (a) Explain what is instrumental variable regression

   Ans: Instrumental variable regression provides a way to handle least squares estimation with training data that is noisy on both the input and the output observations. In other words, an instrumental variable is a third variable, Z, used in regression analysis when we have endogenous variables that is variables that are influenced by other variables in the model. Instrumental variables are used to account for unexpected behavior between variables. Using an instrumental variable to identify the hidden (unobserved) correlation allows us to see the true correlation between the explanatory variable and response variable, Y. Therefore an instrumental variable is a vector that is correlated with true input vector but that is uncorrelated with the observation noise.

   (b) Explain why LSTD can be interpreted as instrumental variables regression.

   Ans: In LSTD we use a linear function approximator where we address the problem of finding a parameter θ^* which will allow us to compute the value function at state s such that

$$v(s) = \theta^{*T}\phi(s)$$

   where θ^T is the transpose of the vector θ.
   Therefore the value function $V(s)$ can be written as

$$\begin{aligned} V(s) &= \sum_{y \in X} P(x,y)[R(x,y) + \gamma V(y)] \\ = \sum_{y \in X} P(x,y)R(x,y) &+ \gamma \sum_{y \in X} P(s,y)V(y) \\ \bar{r}_x &+ \gamma \sum_{y \in X} P(x,y)\phi_y'\theta^* \end{aligned}$$

   where $\bar{r}_x$ is the expected reward. The expression above can be written as the following

$$\bar{r}_x = \phi_x - \gamma \sum_{y \in X} P(x,y)\phi_y)'\theta^*$$

Thus the scalar output $\bar{r}_x$ is the inner product of the input vector $\phi_x - \gamma \sum_{y \in X} P(x, y)\phi_y)$ and θ^*

$$r_t = (\phi - \gamma \sum_{y \in X} P(x_t, y)\phi_y)'\theta^* + (r_t - \bar{r_t})$$

where the noise $(r_t - \bar{r_t})$ is uncorrelated to the input vector $w_t = (\phi - \gamma \sum_{y \in X} P(x_t, y)\phi_y)$ Thus the term above can be written as
$\hat{w}_t = \phi_t - \gamma\phi_{t+1}$ since we do not know the transition function or the state of the markov chain at each time step. Hence it reduces to the following equation

$$r_t = (\phi_t - \gamma\phi_{t+1})'\theta^* - (\gamma P(x_t, y)\phi_y - \gamma\phi_{t+1})'\theta^* + (r_t - \bar{r_t})$$

From the equation above it can be seen that the ϕ is the instrumental variable because it is uncorrelated with the input observation noise. Thus the LSTD can be interpreted as the instrumental variable regression.