Cem Daloğlu
CS464-1 HW1

1) Probability Questions

Question 1.1

In order to have one turnover in last eight matches, the team can either win first k matches or lose all the remaining matches or vice-versa. It can be shown in a probabilistic way;

$$P(W = k) * P(L = 8 - k) = (0.6)^k (0.4)^{8-k}$$

$$W = winning\ first\ k\ games,\ \ L = losing\ \ last\ 8 - k\ games,\ \ 0 < k < 8, k \in \mathbb{Z}^+$$

So, the probability for having exactly 1 turnover for the last eight games is,

$$P(1\ turnover) =$$

$$2\big(P(W = 1) * P(L = 7) + P(W = 2) * P(L = 6) + \cdots + P(W = 7) * P(L = 1)\big)$$

$$= 0.063252$$

The equation is multiplied by two because same probabilities can be applied to losing first k games and winning last 8-k games case.

Question 1.2

If the player A wins after the first draw condition, its probability equals to;

$$P(AA) = p^2$$

$$A{:}\, player\ A\ wins,\ \ B{:}\, player\ B\ wins$$

Player A can also win after the second draw conditions appeared whose probability equals to;

$$\big(P(AB) + P(BA)\big)P(AA) = (p(1-p) + (1-p)p)p^2$$

$$= 2p(1-p)p^2$$

So, with probability $2p(1-p)$, the game returns to its initial position and the player A can win after n draw conditions appeared.

$$P(Player\ A\ wins) = P(AA) + \big(P(AB) + P(BA)\big)P(AA) + \big(P(AB) + P(BA)\big)^2 P(AA) + \cdots$$

$$= P(AA) \sum_{n=0}^{\infty} \big(P(AB) + P(BA)\big)^n$$

$$= p^2 \sum_{n=0}^{\infty} \big(2p(1-p)\big)^n$$

Also $p \in (0,1)$, then $0 < 2p(1-p) < 1$. From the sum of a geometric series, the equation converges to,

$$P(Player\ A\ wins) = \frac{p^2}{1 - 2p(1 - p)}$$

The following figure is obtained as the probability as a function of p for $p \in [0,1]$;

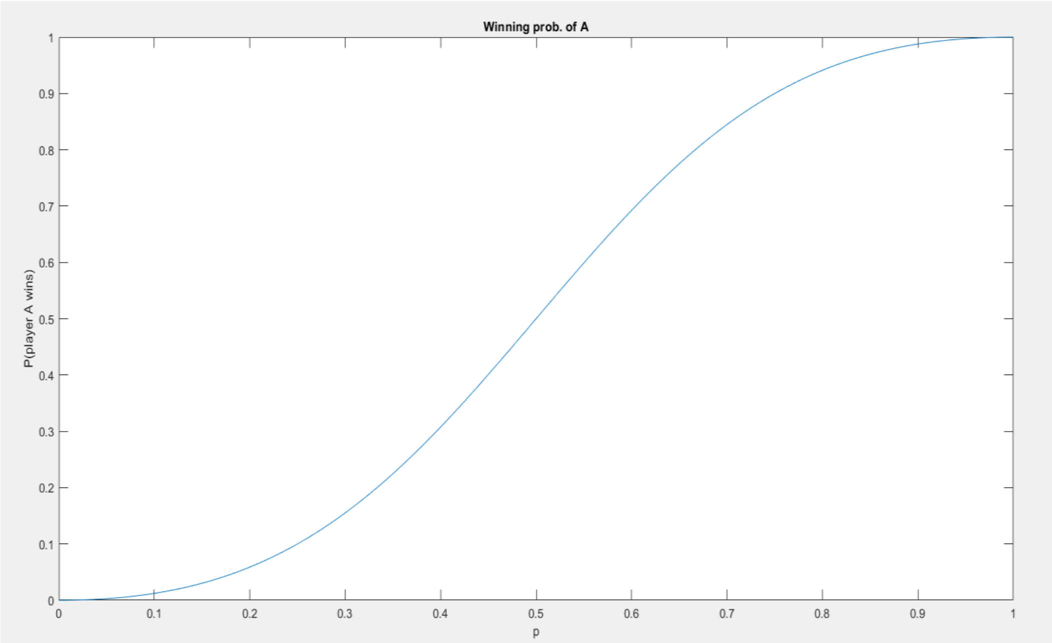

Figure 1: Probability of Player A winning

2) Principal Component Analysis

Question 4.1

For PCA we want to find the directions of the largest variance so we want to find a unit vector ω_1 that maximizes the variance $Var(z_1)$. By maximizing the first principal component, the eigenvalue of the first component will be in the direction of the maximum variance. It is a constrained optimization problem and it can be solved by Lagrange multiplier method. The following Lagrange equations are obtained;

$$\mathcal{L}(\omega, \lambda) \equiv \sigma_\omega^2 - \lambda(\omega^T \omega - 1) \ \ (1)$$

$$\sigma_\omega^2 = \omega_1^T \Sigma \omega_1 \ \ (2)$$

$$\frac{\partial \mathcal{L}}{\partial \omega_1} = 2\Sigma\omega_1 - 2\lambda\omega_1 = 0 \ \ (3)$$

The sign of λ can be positive or negative so, the equation (1) should be solved for the stationary point of $\mathcal{L}(\omega, \lambda)$ which is done in the equation (3).

To clarify, ω_1 is an eigenvector, λ is an eigenvalue of covariance matrix Σ. In order to maximize the variance, $\Sigma\omega_1$ is replaced with $\lambda\omega_1$ in equation (2) we get;

$$max(\omega_1^T \Sigma \omega_1) = max(\lambda \omega_1^T \omega_1)$$

Since eigenvectors are normalized $\omega_1^T \omega_1 = 1$;

$$max(\lambda \omega_1^T \omega_1) = max(\lambda)$$

In other words, the first principal component of the system comes from the eigenvector that gives the largest variance. Therefore, to maximize the variance with respect to the eigenvector, the maximizing vector will be the one with the largest eigenvalue.

Question 4.2

Similar as the previous problem, Lagrange multiplier will be used to prove the question statement. Lagrange equation for the second principal component is;

$$\mathcal{L}(\omega, \lambda_2, \phi) \equiv \omega_2^T \Sigma \omega_2 - \lambda_2(\omega_2^T \omega_2 - 1) - \phi(\omega_2^T \omega_1 - 1) \quad (1)$$

$$\frac{\partial \mathcal{L}(\omega, \lambda_2, \phi)}{\partial \omega_2} = 2\Sigma \omega_2 - 2\lambda_2 \omega_2 - \phi \omega_1 = 0 \quad (2)$$

By multiplying the equation (2) with ω_1^T;

$$2\omega_1^T \Sigma \omega_2 - 2\lambda_2 \omega_1^T \omega_2 - \phi \omega_1^T \omega_1 = 0 \quad (3)$$

We know that, $\omega_1^T \omega_1 = 1$ and $\omega_1^T \omega_2 = 0$.

$$0 - 0 - \phi = 0$$

Thus, $\phi = 0$. So, the equation (3) become;

$$\Sigma \omega_2 - \lambda_2 \omega_2 = 0$$

$$\Sigma \omega_2 = \lambda_2 \omega_2$$

When the variance is maximized with respect to the second principle component with these two constraints, ω_2 will be the eigenvector with the second largest eigenvalue of the system.