Homework 1

Giulio Nenna (s292399@studenti.polito.it)

December 11, 2021

Realizzato in collaborazione con Alessandro Bonaduce (s289906@studenti.polito.it), Davide Grande (s292174@studenti.polito.it), Ciro Balsamo (s289363@studenti.polito.it)

Esercizio 1.

(a) Per la dinamica P (non lazy) distinguiamo 2 casi:

- Se $\frac{n}{2}$ è dispari allora il grafo **non** è aperiodico, pertanto la convergenza della dinamica P non è garantita.
- Se $\frac{n}{2}$ è pari allora siano

$$p_1 = \{0, 1, 2, \dots, n-1, 0\}$$

$$p_2 = \{0, 1, 2, \dots \frac{n}{2}, 0\}$$

due cicli di lunghezza rispettivamente n e $\frac{n}{2} + 1$.

$$\frac{n}{2} \quad \text{pari} \quad \implies n = 4k,\, k \in \mathbb{N}$$

Abbiamo quindi trovato due cicli di lunghezza $4k$ e $2k+1$ che sono numeri coprimi tra loro, pertanto il grafo è aperiodico e la convergenza è garantita. In particolare

$$x(t) \xrightarrow{t \to \infty} (\pi' x(0))\mathbf{1}$$

Poiché il grafo R_n è regolare, allora $\pi = \frac{1}{n}\mathbf{1}$. La dinamica Q converge sempre dal momento che vengno aggiunti self-loop ad ogni nodo e quindi il grafo risulta sempre aperiodico. In particolare, poiché:

$$\pi' P = \pi' \iff \pi' Q = \pi'$$

a parità di condizioni iniziali $x(0)$ la dinamica Q converge allo stesso consenso della dinamica P (nel caso in cui $\frac{n}{2}$ sia pari)

(b)

$$P = \frac{1}{3}W \implies Q = \frac{1}{6}W + \frac{1}{2}I$$

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \cdots & \frac{1}{6} & \cdots & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} & \cdots & \frac{1}{6} & \cdots \\ \vdots & & & & & \end{bmatrix}$$

Q è una matrice circolare la cui prima riga q è così composta:

$$q_0 = \frac{1}{2}, \quad q_1 = q_{\frac{n}{2}} = q_{n-1} = \frac{1}{6}, \quad q_k = 0\,\forall k \notin \{1, 2, \frac{n}{2}, n-1\}.$$

(gli indici sono shiftati di una posizione, da 0 a $n-1$). Poiché Q è circolare allora lo spettro sarà dato da:

$$\lambda_k = \sum_{j=0}^{n-1} q_j \omega_k^j, \quad k = 0, \ldots, n-1$$

$$\omega_k = \exp\left[\frac{2\pi i}{n}k\right].$$

Pertanto [1]:

$$\begin{aligned} \lambda_k &= \frac{1}{2}\omega_k^0 + \frac{1}{6}\omega_k^1 + \frac{1}{6}\omega_k^{\frac{n}{2}} + \frac{1}{6}\omega_k^{n-1} = \\ &= \frac{1}{2} + \frac{1}{6}\exp\left[\frac{2\pi i}{n}k\right] + \frac{1}{6}\exp\left[-\frac{2\pi i}{n}k\right] + \frac{1}{6}\exp\left[\pi i k\right] = \\ &= \frac{1}{2} + \frac{1}{3}\cos\left(\frac{2\pi}{n}k\right) + \frac{1}{6}\exp\left[\pi i k\right] \end{aligned}$$

$$\implies \lambda_1 = \frac{2}{3} + \frac{1}{3}\cos\left(\frac{2\pi}{n}\right) = \frac{2}{3} + \frac{1}{3}\left(1 - \frac{2\pi^2}{n^2} + o\left(\frac{1}{n^3}\right)\right) \quad n \to \infty.$$

Quindi:

$$\tau = \frac{1}{1-\lambda_1} \approx \frac{1}{\frac{1}{3} + \frac{2\pi^2}{3n^2}}$$

Ossia per $n \to \infty$, $\tau \to 3$.

[1] Stiamo considerando gli indici shiftati di una posizione per cui λ_1 sarà il secondo autovalore di Q.

(c) Il grafo barbell è aperiodico ed è costituito da due "cluster" completi collegati da un unico edge. Se $\lambda_1 \geq \lambda_2, \ldots$ sono gli autovalori della relativa matrice P, allora vale la seguente:

$$\frac{1}{2}\phi_{\mathcal{G}}^2 \leq 1 - \lambda_2 \leq 2\phi_{\mathcal{G}}.$$

Poiché $Q = \frac{1}{2}P + \frac{1}{2}I$, allora, se $\lambda_2^{(Q)}$ è il secondo autovalore di Q, abbiamo:

$$\lambda_2^{(Q)} = \frac{1}{2}\lambda_2 + \frac{1}{2}.$$

Pertanto vale:

$$\frac{1}{4}\phi_{\mathcal{G}}^2 \leq 1 - \lambda_2^{(Q)} \leq \phi_{\mathcal{G}} \implies \frac{1}{\phi_{\mathcal{G}}} \leq \tau \leq \frac{4}{\phi_{\mathcal{G}}^2}$$

dove τ è il tempo di rilassamento della dinamica lazy nel grafo barbell. Calcoliamo quindi $\phi_{\mathcal{G}}$:

$$\phi_{\mathcal{G}} = \min_{\mathcal{U} \subseteq \mathcal{V}} \phi(\mathcal{U}) \quad \text{s.t.} \quad 0 \leq w_{\mathcal{U}} \leq \frac{1}{2}\mathbb{1}^T w$$

dove

$$w_{\mathcal{U}} = \sum_{i \in \mathcal{U}} w_i \quad \text{e} \quad \phi(\mathcal{U}) = \frac{\sum\limits_{i \in \mathcal{U}} \sum\limits_{j \in \mathcal{U}^C} W_{i,j}}{w_{\mathcal{U}}}.$$

Consideriamo ora il taglio lungo l'edge che separa i due "cluster" del grafo. Otteniamo:

$$\sum_{i \in \mathcal{U}} \sum_{j \in \mathcal{U}^C} W_{i,j} = 1.$$

Questo vale perchè la matrice di adiacenza di u grafo barbell è composta nel seguente modo:

$$W = \begin{bmatrix} \mathbb{1}\mathbb{1}^T - I & F \\ F^T & \mathbb{1}\mathbb{1}^T - I \end{bmatrix}$$

dove

$$F \in \mathbb{R}^{\frac{n}{2} \times \frac{n}{2}},\ (F)_{\frac{n}{2},1} = 1,\ (F)_{i,j} = 0 \quad \text{altrimenti}$$

Scegliendo il taglio sull'edge che separa i due cluster, la quantità $\sum\limits_{i \in \mathcal{U}} \sum\limits_{j \in \mathcal{U}^C} W_{i,j}$ corrisponde a sommare tutti gli elementi della matrice

F che ha un unico elemento diverso da 0 ed è pari a 1. Abbiamo inoltre che:

$$w_{\mathcal{U}} = \left(\frac{n}{2} - 1\right)\frac{n}{2} + 1$$

Perché tutti i nodi in $\mathcal{U}$ hanno grado $\frac{n}{2} - 1$ tranne uno che è quello collegato con l'altro cluster. Quest'ultimo ha grado $\frac{n}{2}$. In totale ci sono $\frac{n}{2}$ nodi in $\mathcal{U}$. Questo taglio è quello con la *bottleneck ratio* più alta, pertanto:

$$\phi_{\mathcal{G}} = \frac{1}{\frac{n^2}{4} - \frac{n}{2} + 1}$$

da cui:

$$O(n^2) \leq \tau \leq O(n^4)$$

(d) Scegliamo come condizioni iniziali per la convergenza al consenso le seguenti:

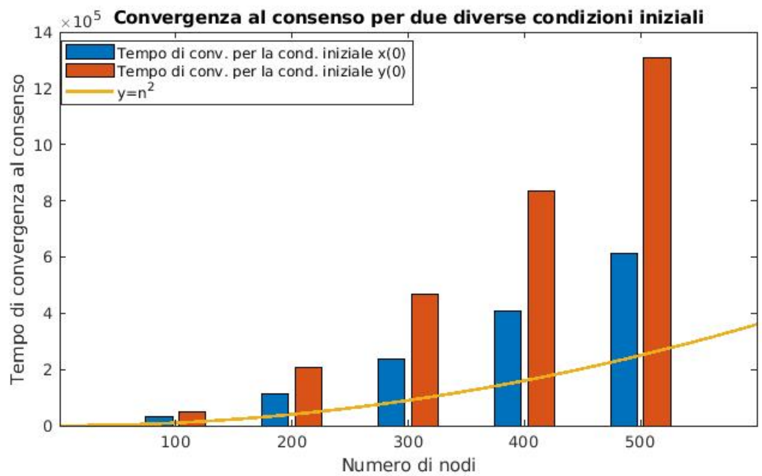
$$x(0) = (0, 1, 0, 1, 0, 1, \dots)$$

$$y(0) = (0, 0, 0, \dots, 0, 1, 1, 1, \dots, 1)$$

Nel caso di $x(0)$, **in ciascuno dei due cluster del grafo barbell,** metà dei nodi ha opinione 1 mentre l'altra metà ha opinione 0. Nel caso di $y(0)$ tutti i nodi di un cluster hanno opinione 1 mentre tutti i nodi dell'altro cluster hanno opinione 0.

```
clear
close all
clc


for i=1:5
n = 100*i;

C=zeros(n/2, n/2);
    C(n/2, 1) = 1;
    Wc=ones(n/2, n/2);
    Wc=Wc-eye(n/2);
    W=[Wc, C;
       C',Wc];
    G=graph(W);
    deegre_centrality = G.degree';
    P = diag(deegre_centrality.^-1)*W;
    Q = 0.5*P + 0.5*eye(n);

    e=ones(n,1);
```

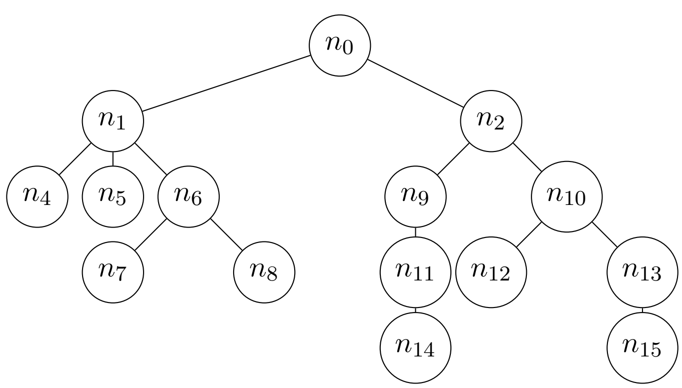
```
     x_0 = zeros(n,1);
     y_0 = zeros(n,1);
     x_0(1:2:n)=e(1:2:n); %convergenza veloce
     y_0(1:n/2)=e(1:n/2); %convergenza lenta



    %---SIMULAZIONE DELLA DINAMICA---------------
     toll=10^-8;
    %--simulazione convergenza veloce------------
    x=x_0;
    x_new=Q*x;
    x_counter=1;
    x_consensus = 0.5*ones(n,1);
    while norm(x_consensus-x_new)>toll
        x=x_new;
        x_new = Q*x;
        x_counter=x_counter+1;
    end
    x_counter_seq(i)=x_counter;

   %--simulazione convergenza lenta------------
    y=y_0;
    y_new=Q*y;
    y_counter=1;
    while norm(x_consensus-y_new)>toll
        y=y_new;
        y_new = Q*y;
        y_counter=y_counter+1;
    end
    y_counter_seq(i)=y_counter;
end

%--plot----------------------------------------
x=linspace(0,600);
figure(1)
bar(100:100:500, [x_counter_seq',y_counter_seq']); hold on
plot(x, x.^2, 'linewidth', 2);
title('Convergenza al consenso per due diverse condizioni iniziali')
legend('Tempo di conv. per la cond. iniziale x(0)','Tempo di conv. per la cond. iniziale y(0)', 'y=n^2');
xlabel('Numero di nodi');
ylabel('Tempo di convergenza al consenso')
```



Come si può notare la convergenza per la condizione $y(0)$ scala molto più lentamente della condizione $x(0)$ rispetto al numero di nodi. Entrambe sono tuttavia sopra il limite inferiore teorico calcolato al punto precedente. Questo fenomeno dipende dal fatto che nel caso di $y(0)$ il consenso viene raggiunto utilizzando come unico canale di comunicazione il singolo edge che collega i due cluster. Nel caso di $x(0)$ questo non è necessario dal momento che le due opinioni sono distribuite uniformemente lungo tutto il grafo.

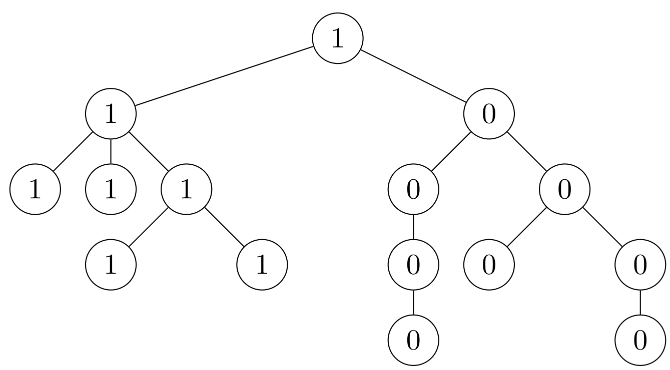
Esercizio 2.

(a)



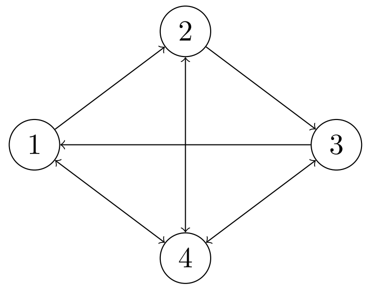
Posizionando il nodo stubborn sul nodo n_0 in figura (ovvero il nodo madre dell'albero) l'agente A avrebbe la certezza che l'opinione finale di ogni agente sia esattamente 1. Il nodo stubborn è infatti global reachable e ciò implica che l'opinione finale di tutti i nodi sarà una combinazione convessa del solo valore 1. A questo punto il giocatore B deve posizionare il suo nodo stubborn con valore 0 in modo da rendere più vicina a 0 possibile la media di opinioni finale. Per farlo è sufficiente porre il suo nodo stubborn su n_2 ovvero sul figlio destro del nodo 1. in questo modo l'opinione finale di tutti i nodi della parte destra del grafo sarà pari a 0. Questa è la scelta ottimale in quanto se posizionasse il nodo stubbron su n_1 (figlio sinistro di n_0) si avrebbero tutti i nodi della parte sinistra dell'albero con opinione finale nulla ma essi sono in numero minore rispetto ai nodi nella parte destra e ciò implicherebbe una media finale maggiore. Se infine, il nodo stubborn venisse posizionato su uno qualsiasi degli altri nodi, l'opinione finale dei nodi compresi tra i due stubbron sarebbe una combinazione convessa dei valori di questi ultimi. Anche in questo caso, il giocatore B non riuscirebbe ad ottenere il minimo valor medio di opinione possibile.

(b) Per le osservazioni al punto precedente il giocatore B dovrebbe posizionare il nodo stubborn ancora una volta su n_2, in questo modo la scelta ottimale per il giocatore A sarebbe necessariamente il nodo n_1, ma la media di opinioni finale sarà più vicina a zero per via della maggior presenza di nodi nella parte destra dell'albero.



Esercizio 3.

(a) La comunità di agenti si presenta sotto forma del seguente grafo:



Dal momento che quest'ultima fortemente connesso, la dinamica delle opinioni raggiunge il consenso:

$$x(t) \xrightarrow{t\to\infty} (\pi' x(0))\mathbb{1}$$

Dal momento che consideriamo $x(0)$ come un vettore di variabili aleatorie indipendenti con le varianze date, allora esso avrà una matrice di varianza-covarianza così composta:

$$VarCov(x(0)) := V = \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.2 \end{bmatrix}$$

Pertanto, utilizzando la formula per la varianza di un prodotto scalare di una V.A. con un vettore costante, si ottiene:

$$Var(\pi' x(0)) = \pi' V \pi.$$

Si procede quindi con il calcolo di π. Essendo il grafo fortemente connesso abbiamo:

$$\pi = \frac{\omega}{\mathbb{1}'\omega}$$

pertanto:

$$Var(\pi' x(0)) \sim 0.0519$$

(b) Sia V la matrice di varianza-covarianza di una comunità. La varianza del consenso γ, se raggiunto, sarà data dalla formula:

$$Var(\gamma) = \pi' V \pi$$

dove π è la distribuzione stazionaria del grafo che connette la comunità. Costruire un grafo che minimizza la varianza del consenso (e quindi l'errore di misura nel caso di dinamiche di opinioni) significa trovare una π che risolve il seguente problema di minimo:

$$\begin{aligned} \text{minimize:} \quad & \pi' V \pi \\ \text{subject to:} \quad & \mathbb{1}'\pi = 1 \\ & \pi \geq 0 \end{aligned}$$

Risolvendo il problema senza considerare il constraint di non negatività si ottiene la funzione lagrangiana:

$$\mathcal{L}(\pi, \lambda) = \pi' V \pi + \lambda(\mathbb{1}'\pi - 1)$$

e risolvendo il sistema:

$$\begin{cases} \nabla_\pi \mathcal{L} = 2V\pi + \lambda \mathbb{1} = 0 \\ \nabla_\lambda \mathcal{L} = \mathbb{1}'\pi - 1 = 0 \end{cases}$$

si trova la soluzione:

$$\pi^* = \frac{1}{\mathbb{1}' V \mathbb{1}} V^{-1} \mathbb{1}. \tag{1}$$

Osserviamo che, poiché V in quanto matrice di varianza-covarianza, è semidefinita positiva, necessariamente π^* è un vettore non negativo, pertanto il constraint di non negatività è automaticamente soddisfatto. Osserviamo inoltre che, nel caso in cui V sia una matrice diagonale, la soluzione π^* trovata è tale per cui viene dato più peso ai nodi che hanno meno varianza: questo ha senso dal momento che se si considera la varianza come "errore di misura" allora per minimizzare questo errore al consenso bisogna dare più centralità ai nodi "esperti" ossia quelli con varianza bassa.

(c) Consideriamo il nuovo grafo con i self-loop e definiamo le quantità:

$$\begin{aligned} \overline{W} &= diag(\alpha) + W \\ \overline{\omega} &= \omega + \alpha \\ \overline{D} &= diag(\overline{\omega}) \end{aligned}$$

Dove $\overline{W}$ e $\overline{\omega}$ sono rispettivamente la matrice di adiacenza e il vettore degli out-degree del nuovo grafo. Calcoliamo ora la matrice stocastica $\overline{P}$ del nuovo grafo:

$$\begin{gathered} \overline{P} = \overline{D}^{-1}\overline{W} \left(diag(\alpha) + diag(\omega)\right)^{-1} \left(diag(\alpha) + W\right) = \\ = \left(diag(\alpha) + diag(\omega)\right)^{-1} \left(diag(\alpha) + diag(\omega)P\right) = \\ = \left(diag(\alpha) + diag(\omega)\right)^{-1} diag(\omega)P + \left(diag(\alpha) + diag(\omega)\right)^{-1} diag(\alpha) \end{gathered}$$

Poiché $\frac{\alpha}{\alpha+\omega} = 1 - \frac{\omega}{\alpha+\omega}$, se definiamo il vettore $a = \frac{\omega}{\alpha+\omega}$, allora dai precedenti calcoli si ottiene[2]:

$$\overline{P} = diag(a)P + I - diag(a).$$

Osserviamo a questo punto che se definiamo

$$\overline{\pi} = diag(a)^{-1}\pi,$$

dove π è la distribuzione stazionaria del grafo senza self loop, allora segue che:

$$\begin{aligned}\overline{P}'\overline{\pi} &= (diag(a)P + I - diag(a))'\, diag(a)^{-1}\pi = \\ &= \left(P' diag(a) + I - diag(a)\right) diag(a)^{-1}\pi = \\ &= diag(a)^{-1}\pi - \pi + P'\pi = \overline{\pi}.\end{aligned}$$

Pertanto $\overline{\pi}$ è una misura stazionaria del nuovo grafo. Affinché $\overline{\pi}$ diventi una distribuzione, bisogna normalizzarla. Quindi:

$$\hat{\pi} = \frac{\mathbb{1}'\omega}{\mathbb{1}'\alpha + \mathbb{1}'\omega}\overline{\pi} = \frac{\alpha+\omega}{\mathbb{1}'\alpha + \mathbb{1}'\omega}$$

è effettivamente una distribuzione stazionaria di $\overline{P}$. Utilizzando ora 1 impostiamo l'equazione per trovare il valore ottimale di α:

$$\hat{\pi}(\alpha) = \frac{\alpha+\omega}{\mathbb{1}'\alpha + \mathbb{1}'\omega} = \frac{1}{\mathbb{1}'V^{-1}\mathbb{1}}V^{-1}\mathbb{1}.$$

poiché

$$V^{-1} = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & \frac{10}{3} & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \qquad \mathbb{1}'V^{-1}\mathbb{1} = \frac{70}{3}, \qquad \mathbb{1}'\omega = 9$$

si ottiene il sistema:

$$\begin{cases} \alpha_1 + 2 = \left(\frac{27+3(\alpha_1+\alpha_2+\alpha_3+\alpha_4)}{70}\right) 10 \\ \alpha_2 + 2 = \left(\frac{27+3(\alpha_1+\alpha_2+\alpha_3+\alpha_4)}{70}\right) \frac{10}{3} \\ \alpha_3 + 2 = \left(\frac{27+3(\alpha_1+\alpha_2+\alpha_3+\alpha_4)}{70}\right) 5 \\ \alpha_4 + 3 = \left(\frac{27+3(\alpha_1+\alpha_2+\alpha_3+\alpha_4)}{70}\right) 5 \end{cases}$$

[2] Con questa notazione si intende dire che $a_i = \frac{\omega_i}{\alpha_i+\omega_i}$ pertanto $diag(a) = diag(\frac{\omega}{\alpha+\omega}) = (diag(\alpha) + diag(\omega))^{-1}\, diag(\omega)$.

da cui si ottengono le soluzioni:

$$\alpha^*(h) = \begin{pmatrix} 2h+4 \\ \frac{2}{3}h \\ h+1 \\ h \end{pmatrix}$$

Tra queste, quella che minimizza $\mathbb{1}'\alpha$, è

$$\alpha^* = \begin{pmatrix} 4 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

(d) Analogamente al punto (b) valcoliamo la varianza del consenso come varianza di un prodotto scalare tra un vettore costante e un vettore aleatorio. Questa volta $VarCov(x(0)) = A$, pertanto:

$$Var(\pi' x(0)) = \pi' A \pi \sim 0.1025$$

(e) Utilizzando lo stesso procedimento del punto (c) impostiamo l'equazione per trovare il valore ottimale di α, questa volta con una matrice di varianza covarianza diversa:

$$\hat{\pi}(\alpha) = \frac{\alpha + \omega}{\mathbb{1}'\alpha + \mathbb{1}'\omega} = \frac{1}{\mathbb{1}' A^{-1} \mathbb{1}} A^{-1} \mathbb{1}.$$

poiché

$$\frac{1}{\mathbb{1}' A^{-1} \mathbb{1}} A^{-1} \mathbb{1} = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

si ottiene il sistema:

$$\begin{cases} \alpha_1 + 2 = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 9)\frac{1}{4} \\ \alpha_2 + 2 = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 9)\frac{1}{4} \\ \alpha_3 + 2 = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 9)\frac{1}{4} \\ \alpha_4 + 3 = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 9)\frac{1}{4} \end{cases}$$

da cui si ottengono le soluzioni:

$$\alpha = \begin{pmatrix} 1+h \\ 1+h \\ 1+h \\ h \end{pmatrix}.$$

Tra queste, quella che minimizza la quantità $\mathbb{1}'\alpha$ è pari a:

$$\alpha^* = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$