6.13 Esercizio. Si costruisca il duale di:

$$
\begin{aligned}
\max \quad & \sum_{s\in S}\sum_{d\in D(s)} r(s,d)\, x(s,d) \\
& \sum_{d\in D(s)} x(s,d) = \sum_{t\in S}\sum_{d\in D(t)} p(t,s,d)\, x(t,d) && \forall s\in S \\
& \sum_{s\in S}\sum_{d\in D(s)} x(s,d) = 1 \\
& x(s,d)\ge 0 && \forall d\in D(s),\quad \forall s\in S
\end{aligned}
$$

dove le variabili sono $x(s,d)$. Il problema massimizza il guadagno medio in un processo markoviano di decisione ad orizzonte infinito. S è l'insieme dgli stati, $D(s)$ è l'insieme delle decisioni disponibili nello stato s, $x(s,d)$ è la probabilità che il sistema sia nello stato s e venga presa la decisione d, $r(s,d)$ è il guadagno dovuto alla decisione d presa nello stato s e $p(t,s,d)$ è la probabilità che avvenga una transizione dallo stato t allo stato s se viene adottata la decisione d. ∎

Soluzione. Il problema viene riscritto come:

$$
\begin{aligned}
\max \quad & \sum_{s\in S}\sum_{d\in D(s)} r(s,d)\, x(s,d) \\
& \sum_{d\in D(t)} x(t,d) = \sum_{s\in S}\sum_{d\in D(s)} p(s,t,d)\, x(s,d) && \forall t\in S \\
& \sum_{s\in S}\sum_{d\in D(s)} x(s,d) = 1 \\
& x(s,d)\ge 0 && \forall d\in D(s),\quad \forall s\in S
\end{aligned}
$$

$$
\begin{aligned}
\max \quad & \sum_{s\in S}\sum_{d\in D(s)} r(s,d)\, x(s,d) \\
& \sum_{s\in S}\sum_{d\in D(s)} ([t=s]-p(s,t,d))\, x(s,d) = 0 && \forall t\in S \\
& \sum_{s\in S}\sum_{d\in D(s)} x(s,d) = 1 \\
& x(s,d) \ge 0 && \forall d\in D(s),\quad \forall s\in S
\end{aligned}
$$

Indicando con $h(t)$ e con g le variabili duali, il duale è:

$$
\begin{aligned}
\min \quad & g \\
& \sum_{t\in S}([t=s]-p(s,t,d))\, h(t) + g \ge r(s,d) \quad \forall d\in D(s),\quad \forall s\in S
\end{aligned}
$$

ovvero

$$
\begin{aligned}
\min \quad & g \\
& g \ge r(s,d) - h(s) + \sum_{t\in S} p(s,t,d)\, h(t) \quad \forall d\in D(s),\quad \forall s\in S
\end{aligned}
$$