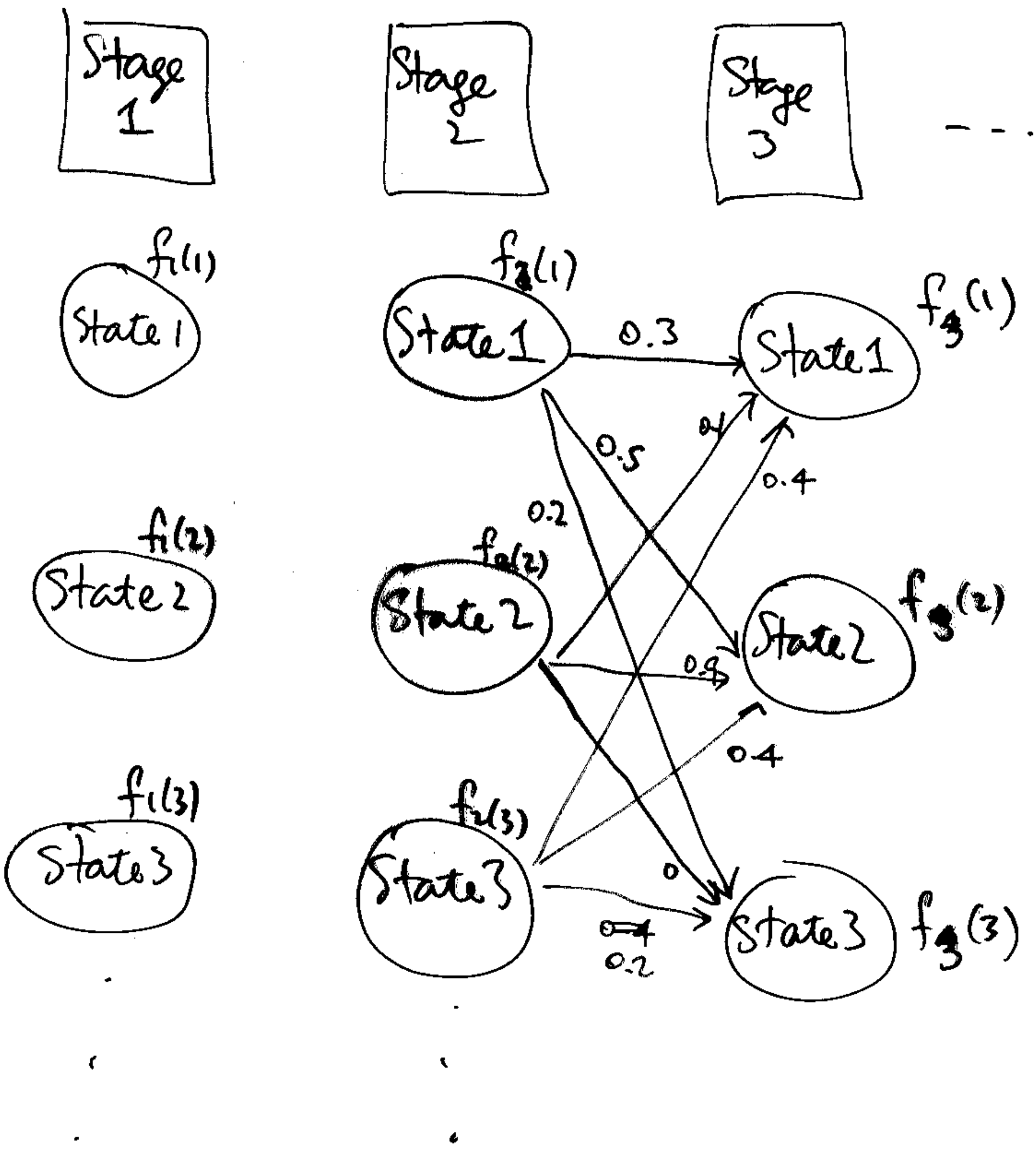
29/10

Probabilistic Dynamic Programming



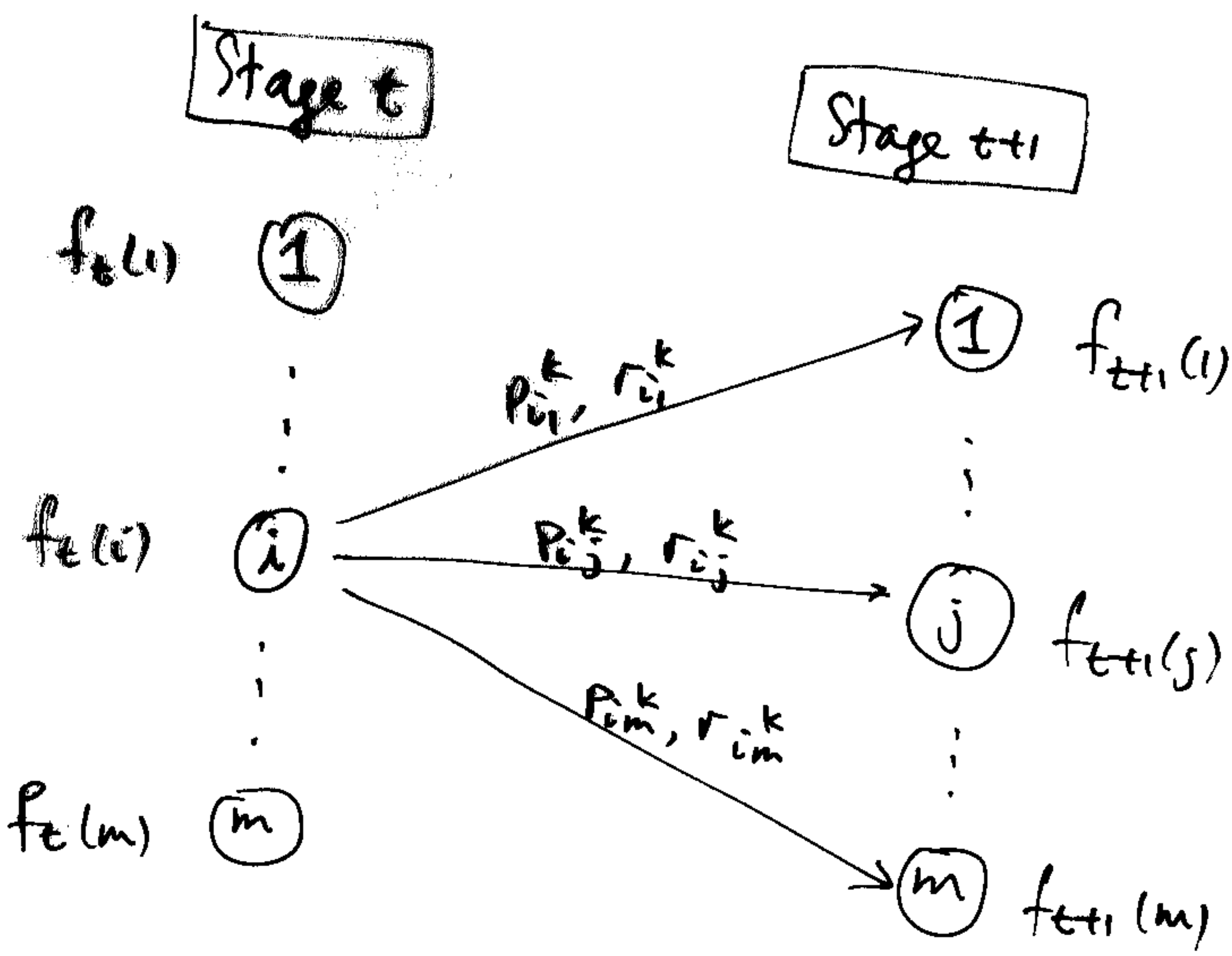
Finite Stage Model

The backward recursive equation relating f_t and f_{t+1}:

$$f_t(i) = \max_k \left\{ r_i^k + \sum_{j=1}^{m} p_{ij}^k f_{t+1}(j) \right\}$$

$$t = 1, \cdots, T, \quad r_i^k = \sum_{j=1}^{m} p_{ij}^k r_{ij}^k$$

for $f_{T+1}(j) \equiv 0 \quad \forall j$.



Infinite Stage Model

$$f(i) = \max_k \left\{ r_i^k + \alpha \sum_{j=1}^{m} p_{ij}^k f(j) \right\}$$

discount factor

$i = 1, \cdots, m$ are the states.

An LP model

To Determine: best decision k whenever state is i

$$\min \; f(1) + f(2) + \cdots + f(m).$$

$$\text{s.t.} \quad f(1) \geq r_1^k + \alpha \sum_{j=1}^{m} p_{ij}^k f(j) \quad \forall k$$

$$\vdots$$

$$f(m) \geq r_m^k + \alpha \sum_{j=1}^{m} p_{ij}^k f(j) \quad \forall k.$$

Optimal sol^n to LP: constraints that are binding correspond to optimal strategies.

A machine in excellent condition earns \$100 profit per week, a machine in good condition earns \$70 per week, and a machine in bad condition earns \$20 per week. At the beginning of any week, a machine may be sent out for repairs ~~~~ at a cost of \$90. A machine that is sent out for repairs returns in excellent condition at the beginning of the next week. If a machine is not repaired, the condition of the machine evolves according to:

This week	Next Week: Excellent	Good	Bad
Excellent	0.7	0.2	0.1
Good	0	0.7	0.3
Bad	0	0.1	0.9

The company wants to maximise its expected discounted profit over an infinite horizon ($\beta = 0.9$).

Use a linear programming model to determine an optimal stationary policy.

Observe that:

$$\boxed{V_i = \max_{d \in \{N,R\}} \left\{ r_{id} + \beta \sum_{j \in \{E,G,B\}} p(j|i,d) V_j \right\} \quad \text{for } i \in E, G, B}$$

An LP model:

$$\min \quad V_E + V_G + V_B$$

s.t.

$$V_E \geq r_{E,NR} + \beta(0.7 V_E + 0.2 V_G + 0.1 V_B) \quad [E,NR]$$

$$V_G \geq r_{G,NR} + \beta(0.7 V_G + 0.3 V_B) \quad [G,NR]$$

$$V_G \geq r_{E,R} + \beta(V_E) \quad [G,R]$$

$$V_B \geq r_{B,NR} + \beta(0.1 V_G + 0.9 V_B) \quad [B,NR]$$

$$V_B \geq r_{B,R} + \beta(V_E) \quad [B,R]$$

$$V_E, V_G, V_B \text{ urs.}$$

Note that $r_{E,NK} = \$100$,

$r_{G,NK} = \$70$, $r_{G,R} = -\$90 + 0 = -\$90$

$r_{B,NK} = \$20$, $r_{B,R} = -\$90$.

LP: $\min V_E + V_G + V_B$

s.t.

$0.37 V_E - 0.18 V_G - 0.09 V_B \geq 100$ — (1)

$0.37 V_G - 0.27 V_B \geq 70$ — (2)

$-0.9 V_E + V_G \geq -90$ — (3)

$-0.09 V_G + 0.19 V_B \geq 20$ — (4)

$-0.9 V_E + V_B \geq -90$ — (5)

V_E, V_G, V_B urs.

$V_E = 668.3$

$V_G = 562.4$

$V_B = 511.5$

Constraints (1), (2), (5) are binding. ∴ optimal strategy: don't replace if machine is in "excellent" or "good" condition. Replace if machine is in "bad" cond^n.