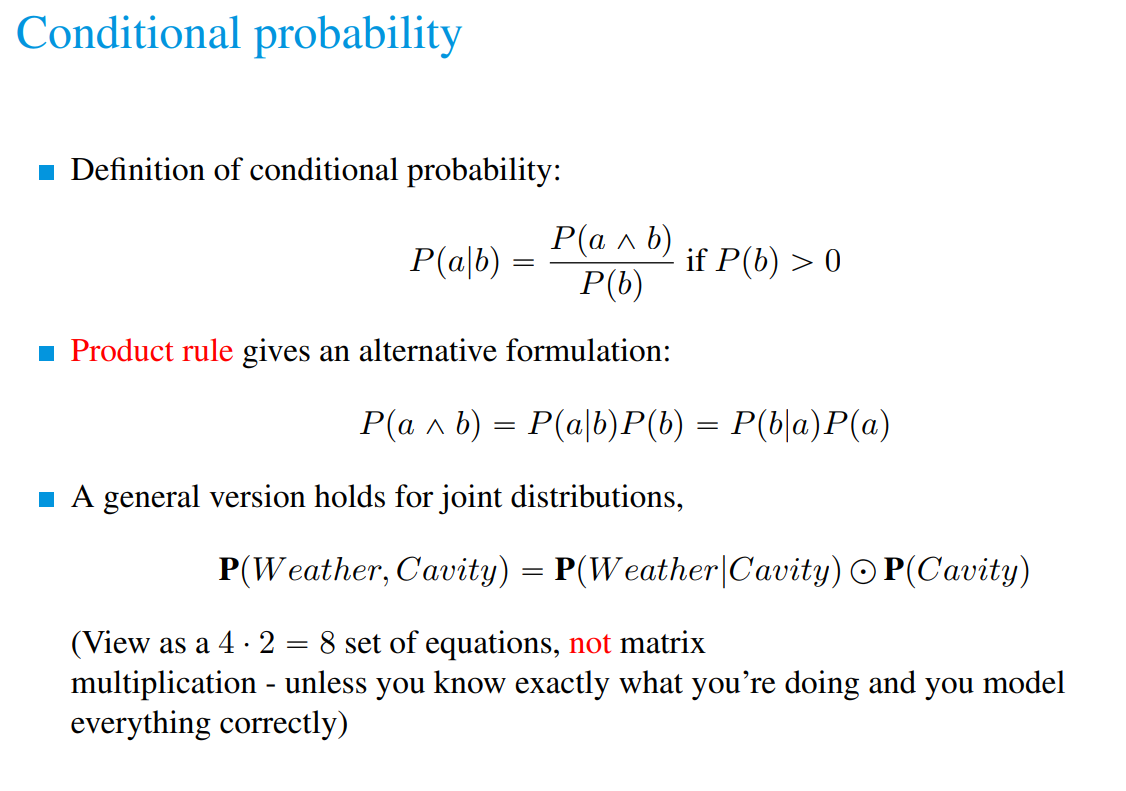
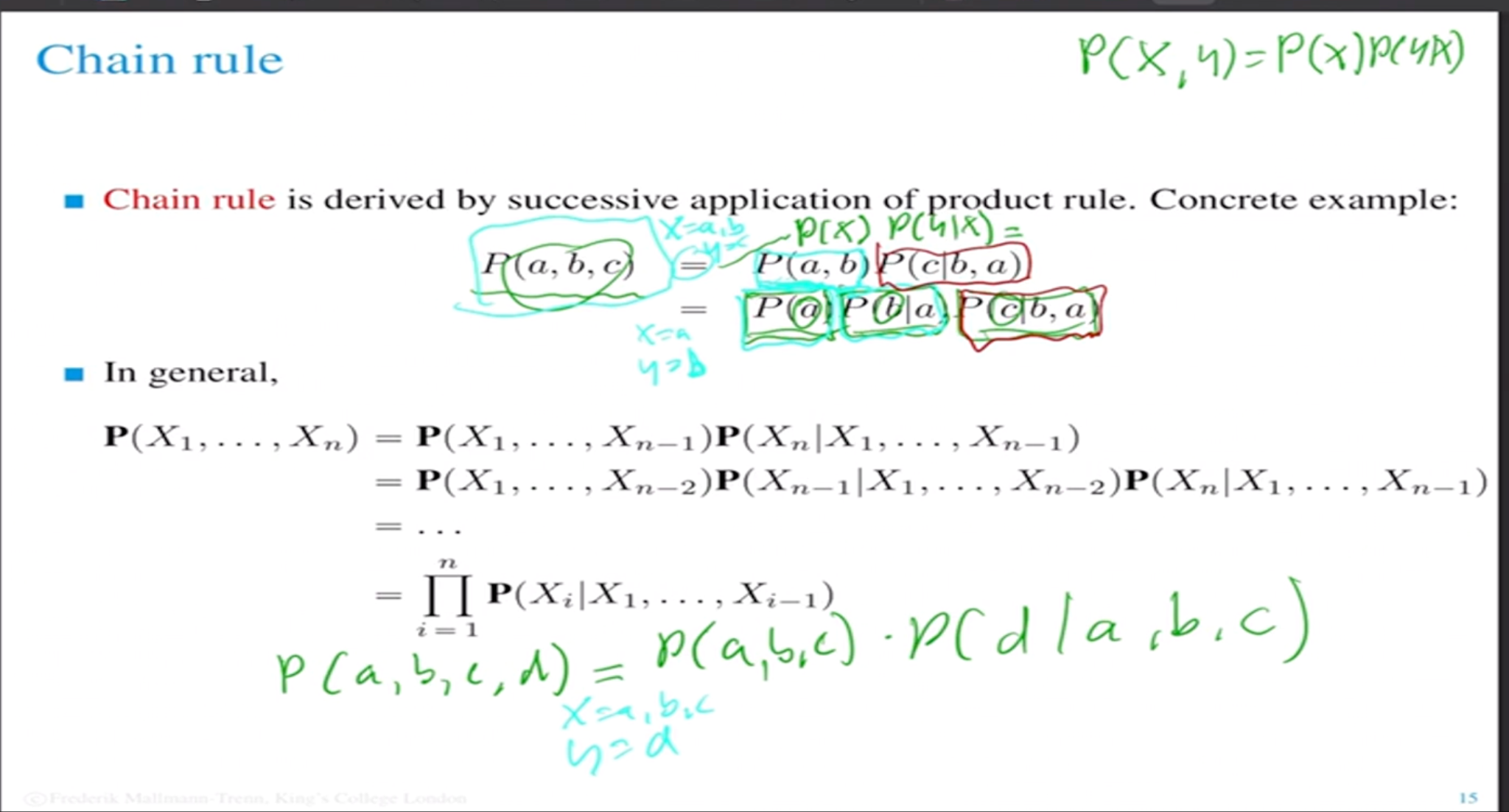
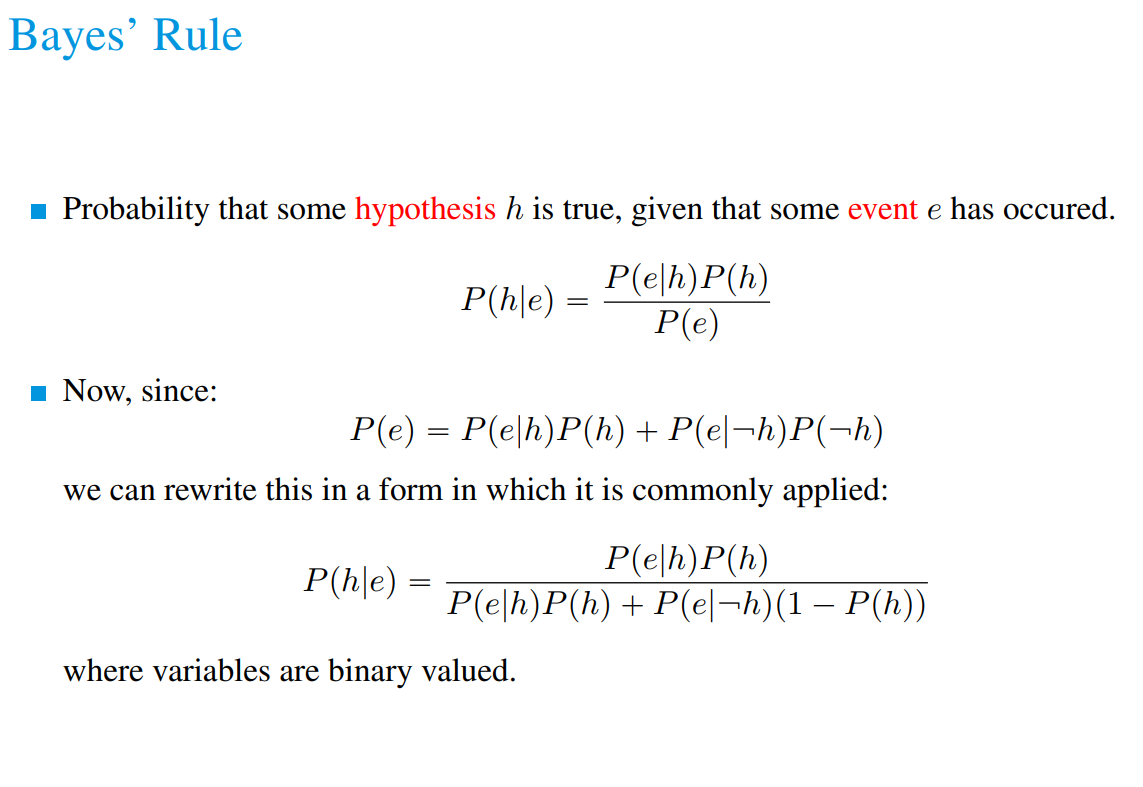
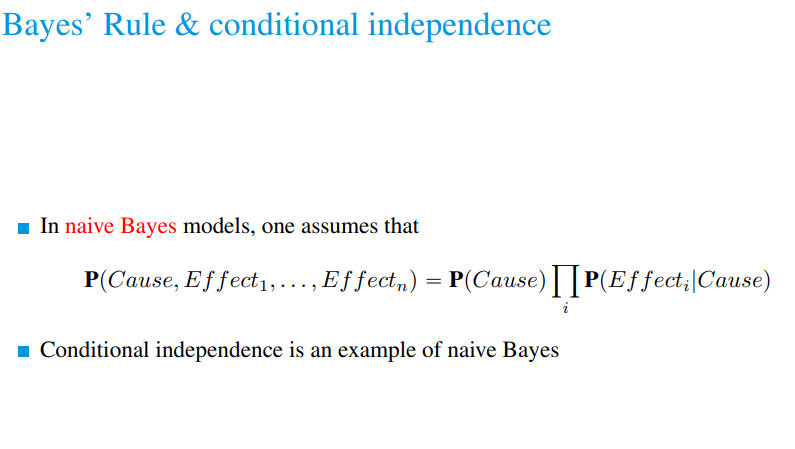
## Week 2

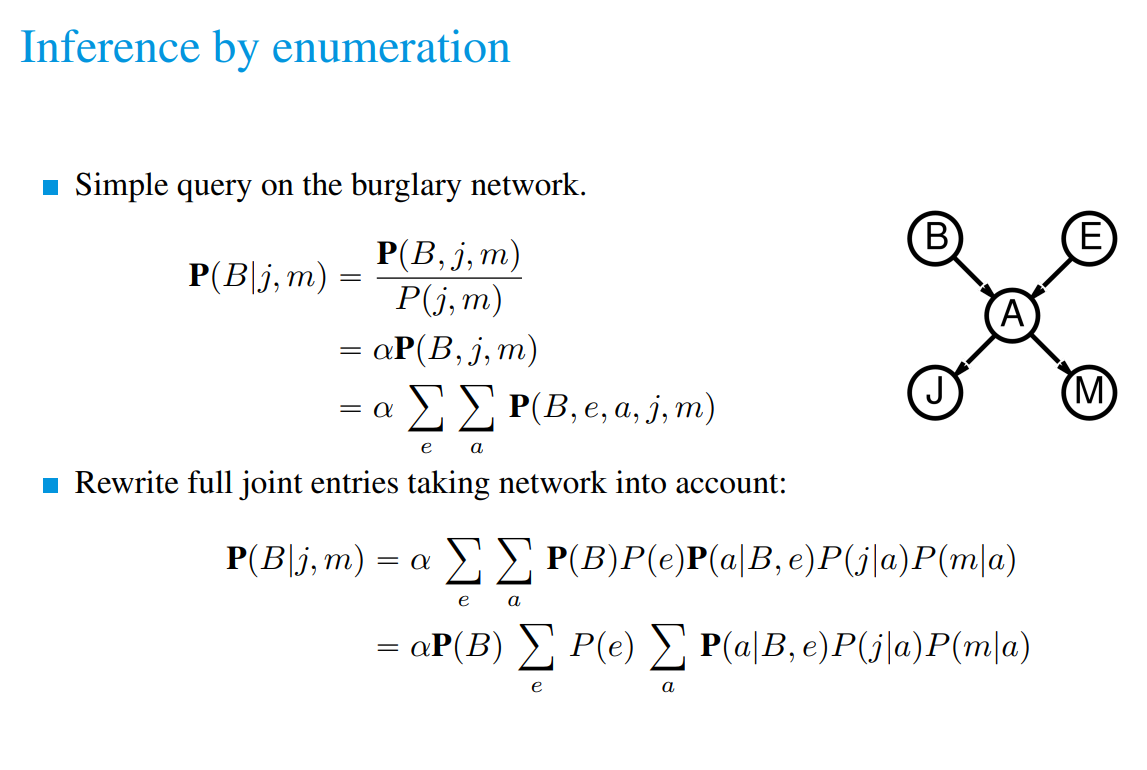




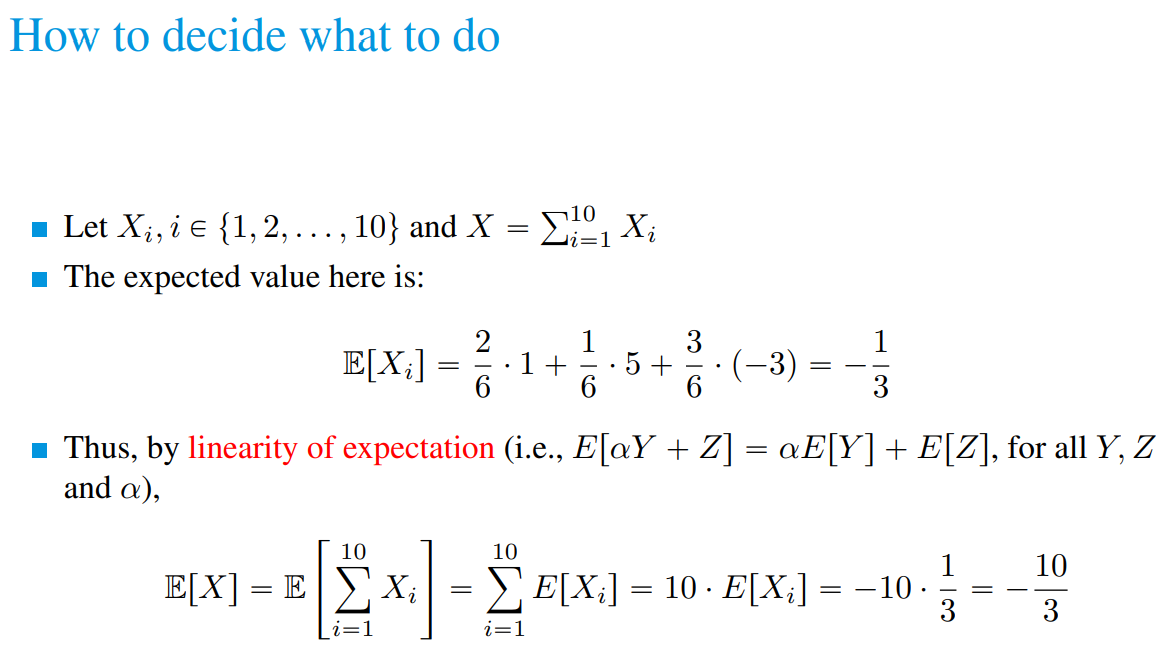


## Week 3

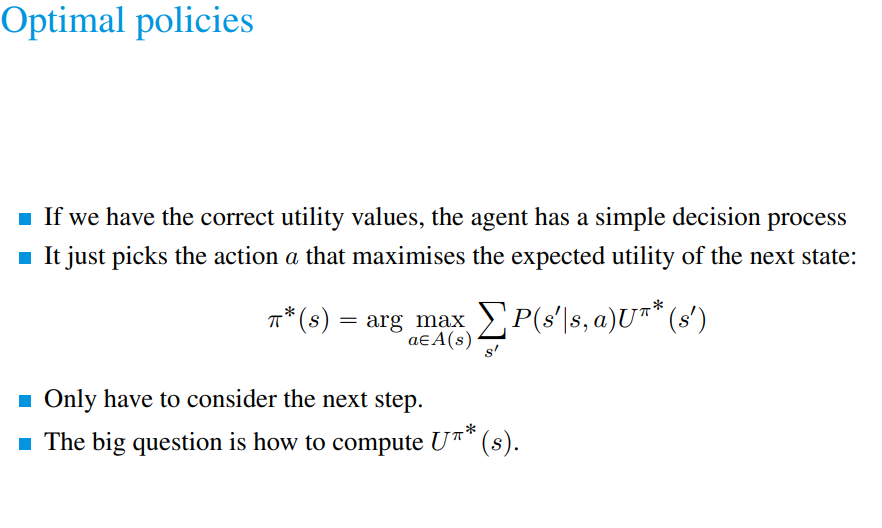


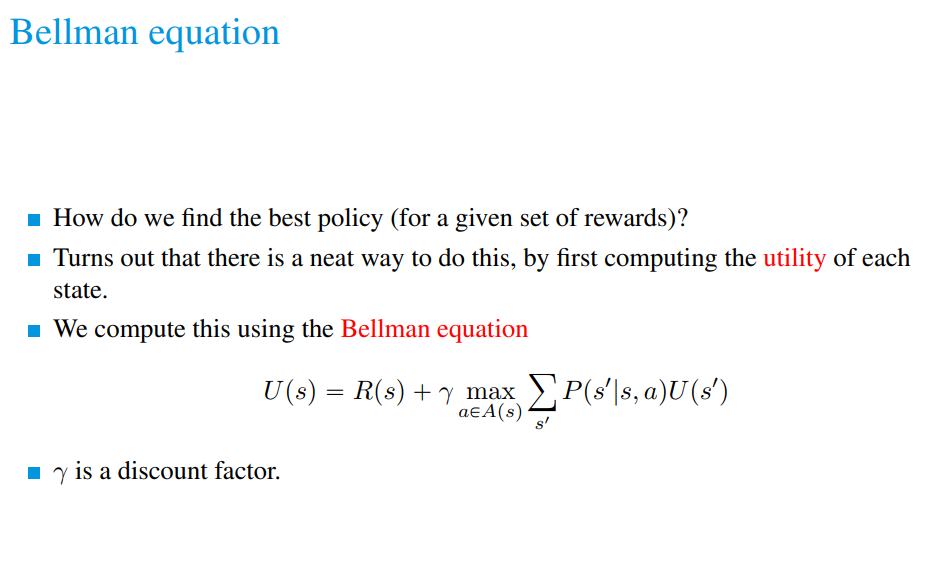


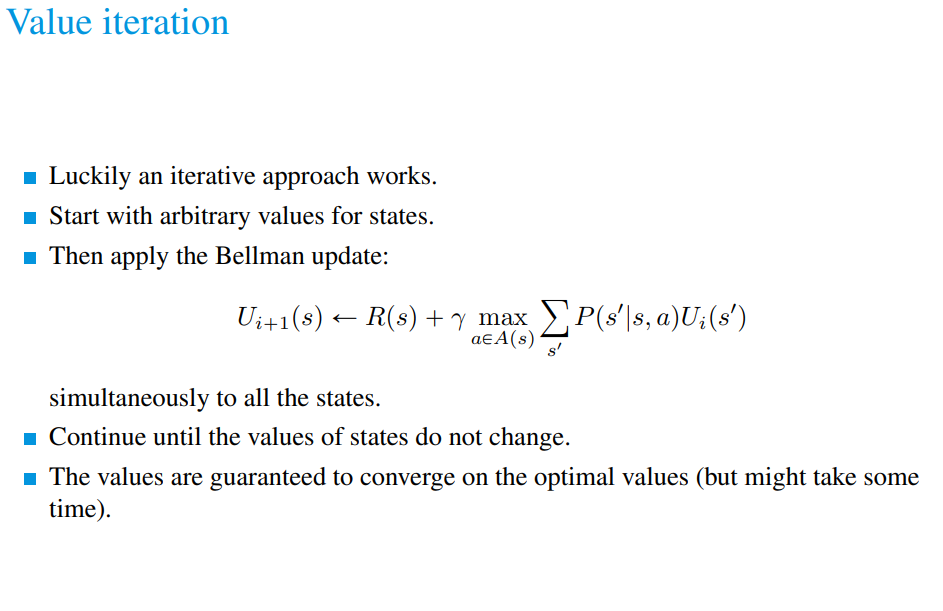
## Week 3

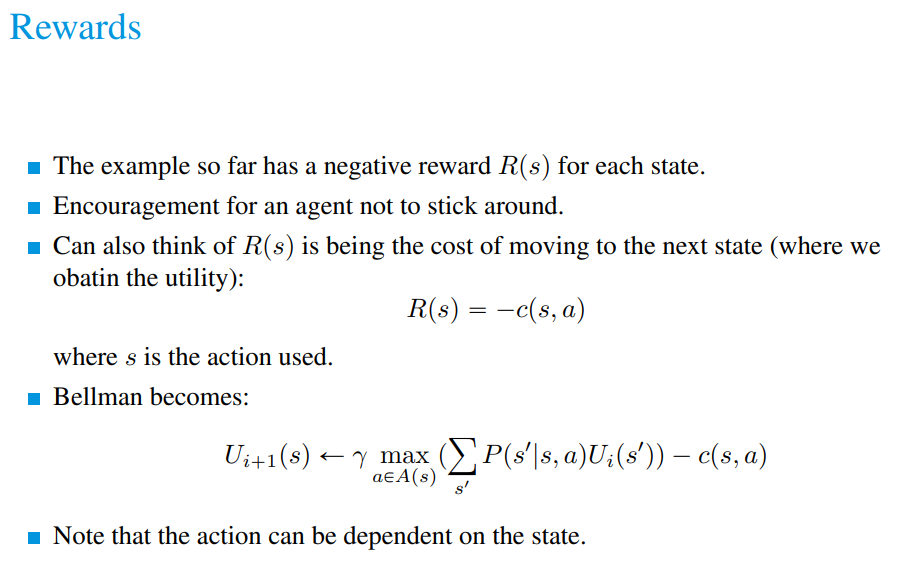


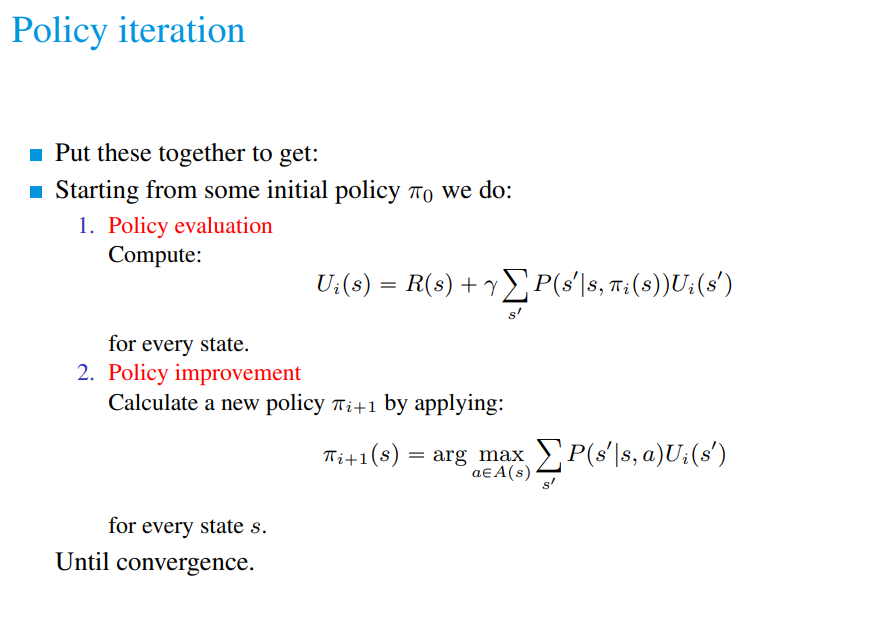
## Week 4





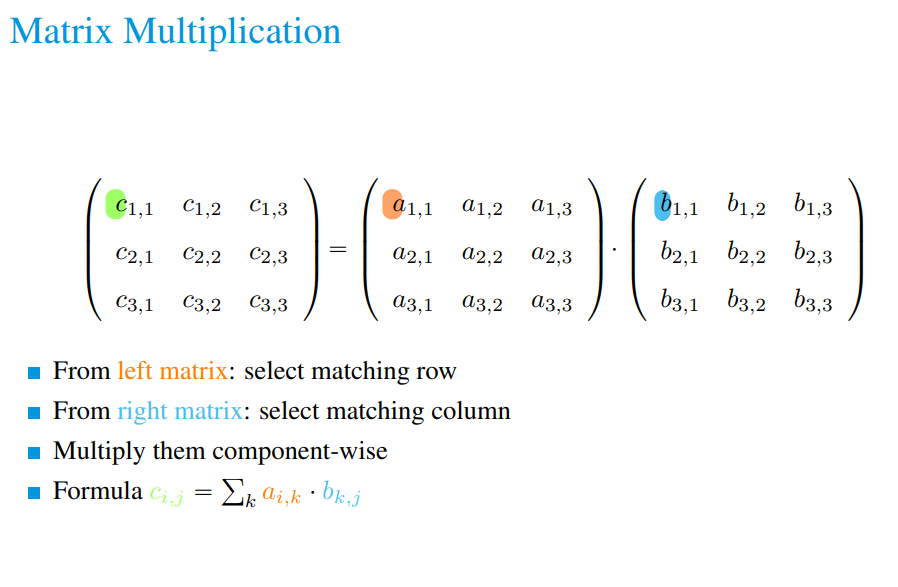


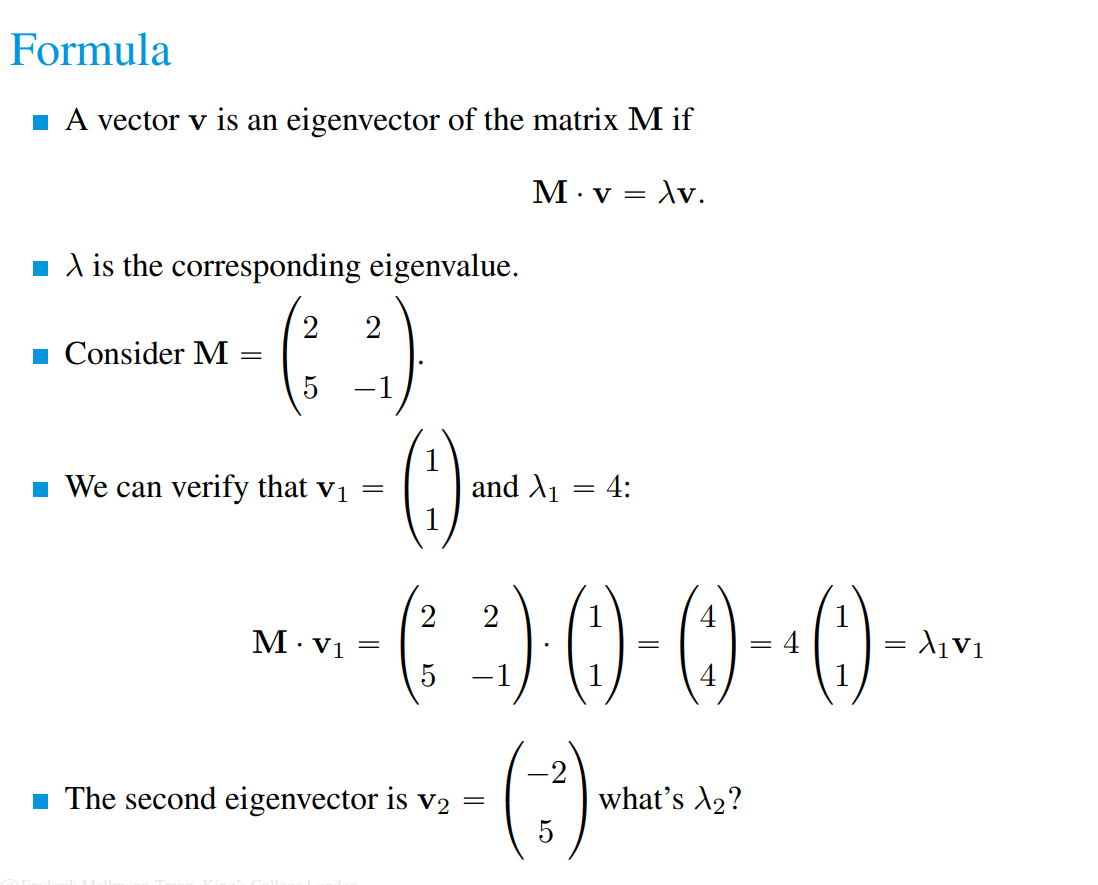


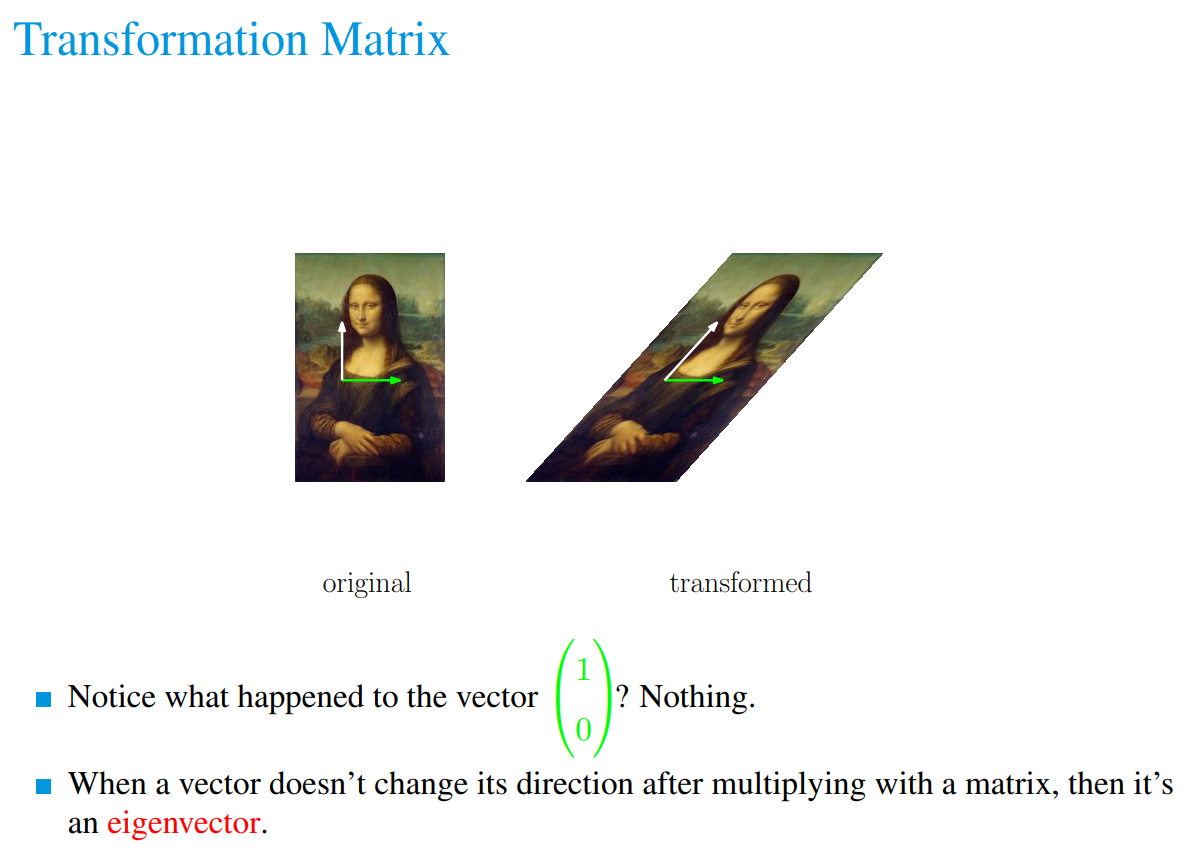


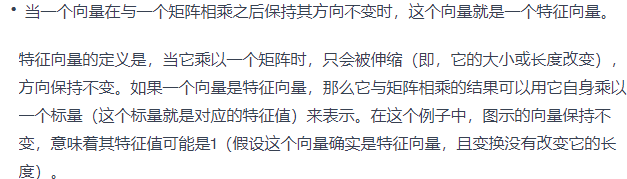
这里的policy evalution之所以没有max在前面是因为policy已经被订好了，只需要对应policy的utility就行，同时因为是提前定好的policy，比如说向左。因此他不是用a来表示，因为a是用来表示一系列actions中的其中一个，这里需要用π来表示。

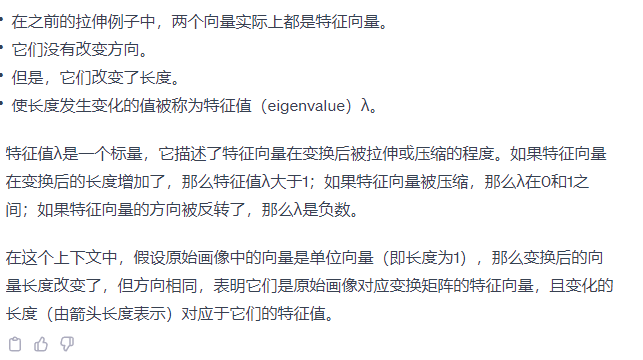
## Week 5





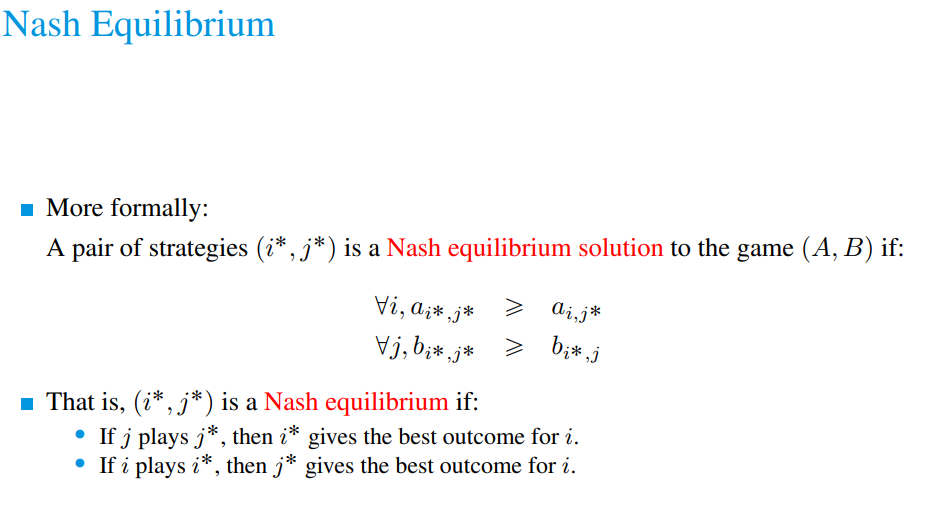




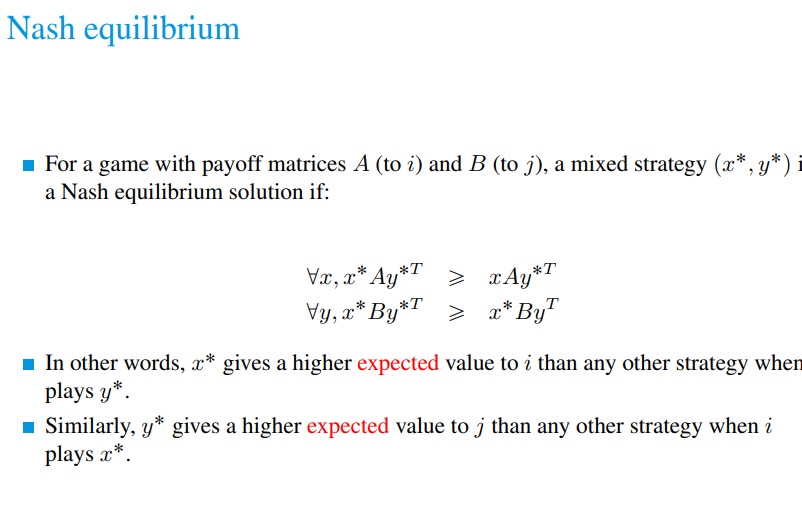


## Week 7

Pure Strategy



Mixed Strategy



## Week 9

