Working algorithm / simulation

(can be more simplified)

Future perspective

Data Strategy

What data do we use

What challenges are there?

How does RL fit with ALM?

ALM as an optimalization problem

RL model for ALM does not compete with the Market. As the bank is the only one aware of its position.



PPO

RL Model



Action Space

Buy or Sell

Tenor

Amount

Needs to be extended to be able to buy or sell multiple swaps at each time step.

Action space is flattened into a discrete value to link to SB3.

Bank Model

Fixed interface: init, step, reset, observation space, action space

Bank Envionment

Observation Space

Zero rates

Projected Cashflows

For a particular position date

Observation space is flattened to a Box (Continuous value between low and high value).

Gym Environment

Bank Model

Generate

NPV / NII / BPV

visualize

Generate Cashflow model

Cashflows

Feature engineering the observation space:

Cashflows are grouped per Month. Zero curves are given as a matrix (rate date, tenor). For simple model no other data is included.

Includes methods:

generate\_mortgage\_contracts

generate\_swap\_contract

generate\_nonmaturing\_deposit

generate\_funding

clear\_swap\_contracts

fixing\_interest\_rate\_swaps

calculate\_npv

calculate\_nii

calculate\_risk

calculate\_bpv

plot\_contracts

plot\_cashflows

reset

step

apply\_action

get\_reward

Interpolate

Zero Rates

Mortgage Rates

visualize

Bank Accounts

Funding

Mortgage Contracts

Forward Rates

Bankmodel II

Introduction

As this model and the steering was becoming increasingly difficult and it became more and more unlikely, I would be able to get a working model ready – I decided to try a more simplified approach. In this model, the actor would still try to steer the interest profile of the bank. But instead of using swaps to modify the interest profile, we will now use bonds to fund the mortgages. By giving the actor the option to directly buy, sell or hold bonds the actor can directly change the interest profile of the bank – and try to match the duration of the cashflows from the mortgages.

The steering would be done to minimize the duration gap between assets and liabilities. In each timestep we will no longer try to calculate the actual BPV Profile – but just measure the absolute difference between assets and liabilities per time bucket.

A cost would be associated with each movement in order to create a situation where the actor needs to decide between minimizing risk and maximizing profit.

As a second step, we may then add the yield curve development, and see if we can create extra profit by not fully matching the duration in certain market circumstances.

The Bank Model

The bank model holds a list of mortgages and associated funding deals. Initially, we generate the mortgages based on a probability distribution in 1, 5, 20, and 30 years mortgages. These mortgages start somewhere in the past but are all still active. Thus simulating an active mortgage portfolio. *Do we really need ‘initial mortgages’ ? This would also mean we need initial funding – and this is something we want the model to find out. Lets remove initial funding.*

The model allows us to buy bonds for a specific tenor. The start date of the bond would be today (so we receive the money directly), and the money would need to be repaid after x years.

At each time step, we increase the position by 1 month. We also add one month of mortgages. The number of new mortgages can fluctuate by +/- 10% but the mean will remain steady over time.

In order to perceive the state of the bank we can calculate the future cashflows. Future cashflows are bucketed on a yearly basis in order to limit the state space in the simulation.

At each iteration, the bank model is reset. The mortgages are initialized and the funding deals are removed.

We calculate the reward at each timestep as a combination of factors. We need to attract funding for the mortgages. Logically you would only fund what you need – so match the funding with the mortgages. We need a little more cash. If we are unable to pay for new mortgages, we will not be able to issue more mortgages. The actor will be ‘encouraged’ to match the duration of the funding with the mortgages by the reward it will receive. So for the reward we punish the model if the funding does not match the mortgages.

In all other cases, the actor will receive a reward for the mortgages outstanding -/- the funding required.

To see how it is going, I would like to monitor the following each step:

* Number of mortgages sold
* Current liquidity
* The mismatch between cashflows (mortgages and funding)

The Bank Environment

The bank environment maps the bank model to the reinforcement learning environment.

Observation space

The observation space will now be the current cashflows – and possibly previous cashflows (to predict movement in the portfolio). We don’t need to know the entire swap curve as rates are not so relevant to the actor.

Action Space

The action space will be multi-discrete. On each time bucket the model can decide to buy, sell or hold. The amount we can sell may be an option. Probably the simplest would be to automatically assign this as a fixed step. We can also automatically assign this as the difference between assets and liabilities.

Mortgage cashflows can be distributed over the next 20 years. (We will not look at interest percentage, prepayment schedule, and duration of the loan). We are just focused on matching the assets with our liabilities. Each timestep will represent a months period, 12 time-steps per year. Cashflows will be bucketed per year for analyses.

The new business would (approximately) need to cover the replacement of the loans that have expired. These new loans will be distributed over the other time buckets –with a fixed duration of 1, 5, 10 or 20 years. To cover these mortgages the agent can buy, hold or sell bonds with a duration of between 2 till 20 years.

Let’s say we can only issue bonds for a fixed amount – for each time bucket - per timestep. As such we are not actually measuring ‘time’ other than time steps (calendar months).

Reward

The actor will need to attract enough funding to fulfil the future cashflow requirements. The goal will be to minimize the combination of the absolute difference in cashflows for each time bucket (the mismatch in the duration). So, a negative reward will be given to these cash flow differences. We can later make this a bit more advanced by only counting it as a negative reward if the mismatch is larger than the risk appetite limit. Additionally, a (fixed) penalty will be incurred directly every time the actor buys or sells a bond.

In the initial model we don’t calculate Net Interest Income, or any other measure of result. So we don’t need to take into account the interest rates of the mortgages and the actual funding cost of the bonds.