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

As this model and the steering was becoming increasingly difficult – 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 zero coupon bonds. By giving the actor the option to directly buy, sell or hold zero coupon bonds the actor can directly change the interest profile of the bank – and try to match the duration of the cashflows from the mortgages.

Furthermore, the mortgages will be modeled as an aggregated cashflow pattern – not as individual 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.

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 2-week period, 24 time-steps per year. Cashflows will be bucketed per year. For every 24th step – the cashflows will shift one-time bucket to the left. Each timestep new mortgages will be issued. 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. Every 24th time step, the cashflows from the bonds are also shifting to the left – and the cashflows on t0 will need to be repaid. 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.

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.