Lecture 12. Evolutionary Modeling and Genetic Algorithms

The purpose of the lecture: to introduce the essence of the problem, to formulate the main provisions and principles, the goals of evolutionary modeling and to give a general concept of genetic algorithms and their capabilities in evolutionary modeling.

Lecture plan:
Introduction
1 Evolutionary modeling
2 Genetic algorithms
Conclusion

Keywords: open system, potential, model adequacy, profit, evolutionary modeling, sets, self-organization, decomposition, activity, vector, relation, system state, entropy, process state, measure, difference, amount of information, control system, objective function, production function, parameter, experience, interval, coefficients, genetic algorithm, evolution, algorithm, maximum, selection, mutation, coordinates, behavior strategy, rating, banking system, union, chromosomes, data structure, information.

Contents of the lecture:

Introduction

The need for a forecast and adequate assessment of the consequences of human activities (especially negative ones) leads to the need to model the dynamics of changes in the main parameters of the system, the dynamics of the interaction of an open system with its environment (resources, potential, conditions, technologies, etc.) with which the exchange is carried out. resources in a hostile, competitive, cooperative or indifferent relationship. It requires a systematic approach, effective methods and criteria for assessing the adequacy of models, which are aimed not only (not so much) at maximizing criteria such as "profit", "profitability", but also at optimizing relations with the environment. If the criteria of the first type are important, for example, for short and medium-term forecasting and tactical administration, then the second type is for medium and long-term forecasting, for strategic administration. At the same time, it is necessary to select and study a fairly complete and informative system of parameters of the system under study and its environment, develop a methodology for introducing measures of information content and proximity of system states. It is important to note that some of the criteria and measures can often conflict with each other.

Many such socio-economic systems can be described from a unified position, by means and methods of a unified theory – evolutionary.

1 Evolutionary modeling

In evolutionary modeling, the process of modeling a complex socio-economic system is reduced to the creation of a model of its evolution or to the search for admissible states of the system, to a procedure (algorithm) for tracking a set of admissible states (trajectories). At the same time, such attributes of biological evolutionary dynamics are actualized (possible socio-economic interpretations of these attributes for evolutionary modeling are given in brackets), for example:

1. *community (corporation, corporate objects, subjects, environment);*
2. *species diversity and distribution in the ecological niche (types of resource allocation, structure of relationships in a given corporation);*
3. *ecological niche (sphere of influence and functioning, evolution in the market, in business);*
4. *fertility and mortality (production and destruction);*
5. *variability (economic environment, resources);*
6. *competitive relationship (market relations);*
7. *memory (the ability to reproduce cycles);*
8. *natural selection (penalties and incentives);*
9. *heredity (production cycles and their background);*
10. *regulation (investment);*
11. *self-organization and the desire of the system in the process of evolution to maximize contact with the environment in order to self-organize, return to the trajectory of sustainable development, and others.*

When studying the evolution of a system, it is necessary to decompose it into subsystems in order to ensure:

1. *effective interaction with the environment;*
2. *optimal exchange of determining material, energy, information, organizational resources with subsystems;*
3. *the evolution of the system under conditions of dynamic change and reordering of goals, structural activity and complexity of the system;*
4. *controllability of the system, identification of the control subsystem and effective connections with the subsystems of the system, feedback.*

Let there be some system S with N subsystems. For each i-th subsystem, we define the vector $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, \ldots, x_{ni}^{(i)})$ of the main parameters (i.e. parameters without which it is impossible to describe and study the functioning of the subsystem in accordance with with the goals and available resources of the system) and the function $s^{(i)} = s(x^{(i)})$, which we will call the activity function or simply the activity of this subsystem.

Example. In business processes, this concept is close to the concept of business activity.

For the entire system, the state vector of the system x and the activity of the system s(x), as well as the concept of the general potential of the system, are determined.

Example. The activity potential can be determined similarly to the biological potential of the population, for example, using the integral of the activity over a given modeling time interval.

These functions reflect the intensity of processes both in subsystems and in the system as a whole.

Three values $s^{(i)}_{max}$, $s^{(i)}_{min}$, $s^{(i)}_{opt}$ – the maximum, minimum and optimal values of the activity of the i-th subsystem, as well as similar values for the entire system (s_{max}, s_{min}, s_{opt}). As an indicator of the economic state, one can also take the ratio of the value of this indicator to its normalized value, and for a comprehensive account of the influence of parameters on the state of the system, one can use analogs of the measure of information proximity, for example, according to K. Shannon.

If an open economic system (process) is given, and H_0, H_1 is the entropy of the system in the initial and final states of the process, then the measure of information is defined as the difference of the form:

$$\Delta H = H_0 - H_1.$$

A decrease in ΔH indicates the approach of the system to a state of static equilibrium (with available resources), and an increase indicates removal. The value ΔH – the amount of information required to move from one level of system organization to another (with $\Delta H > 0$ - higher, with $\Delta H < 0$ – lower organization).

An approach using the measure according to N. Moiseev is also possible. Let some control system be given, about the states of which only some estimates are known - lower s_{min} and upper s_{max}. The objective control function F(s(t), u(t)) is known, where s(t) is the state of the system at time t, and u(t) is a control from a set of admissible controls, and we assume that u_{opt} – some optimal control from the space U, $t_0 < t < T$, $s_{min} \le s \le s_{max}$. Measure of success in decision making:

$$H = \left|\frac{F_{max} - F_{min}}{F_{max} + F_{min}}\right|$$
$$F_{max} = \max F(u_{opt}, s_{max}), F_{min} = \min F(u_{opt}, s_{min}),$$
$$t \in [t_0; T], \quad s \in [s_{min}; s_{max}]$$

An increase in H indicates the success of the system management (the success of the adopted management decision).

The activities of subsystems directly or indirectly interact with the help of the systemic activity s(x), for example, according to a simple scheme of the form

$$\begin{cases} \dfrac{ds(t)}{dt} = \displaystyle\sum_{i=1}^{n} \varphi^{(i)}\left(s; s^{(i)}, t\right), \\ \dfrac{ds^{(i)}(t)}{dt} = \psi^{(i)}\left(s^{(i)}; s, t\right) \end{cases}$$

The functions j(i), y(i) must reflect the evolution of the system, in particular, satisfy the conditions:

1. *periodicity, cyclicity, for example:*

$$\left(\exists\ 0 < T < \infty, \qquad \forall t: \varphi^{(i)}\left(s; s^{(i)}, t\right) = \varphi^{(i)}\left(s; s^{(i)}, t + T\right),\right.$$
$$\left.\psi^{(i)}\left(s; s^{(i)}, t\right) = \psi^{(i)}\left(s; s^{(i)}, t + T\right)\right);$$

2. *attenuation with a decrease in activity, for example:*

$$(s(x) \to 0\ \ \forall i = 1,2, \dots, n) \Rightarrow \left(\varphi^{(i)} \to 0, \psi^{(i)} \to 0\right);$$

3. *equilibrium and stationarity: the choice (definition) of the function $\varphi^{(i)}$, $\psi^{(i)}$ is carried out in such a way that the system has points of equilibrium state, and $s^{(i)}_{opt}$, s_{opt} are reached at stationary points $x^{(i)}_{opt}$, x_{opt} for short time intervals; over long periods of time, the system can (in accordance with the theory of catastrophes) behave chaotically, spontaneously generating regular, ordered, cyclic interactions (deterministic chaos).*

We do not take into account the mutual activities $\psi^{(ij)}\left(s; s^{(i)}, s^{(j)}, t\right)$ of subsystems i and j. As a function $\varphi^{(i)}$, $\psi^{(i)}$, production functions of the Cobb-Douglas type can be effectively used:

$$\varphi(s) = \prod_{i=1}^{n}\left(\frac{s(t) - s_{max}}{s_{opt} - s_{min}}\right)^{\alpha_i(t)}\left(\frac{s_{max} - s(t)}{s_{max} - s_{opt}}\right)^{-\alpha_i(t)\frac{s_{max}-s_{opt}}{s_{opt}-s_{min}}}$$

In such functions, the α_i parameter is important, reflecting the degree of self-regulation, adaptation of the system. As a rule, it needs to be identified.

The functioning of the system satisfies on each time interval $(t; t + \tau)$ constraints of the form

$$\int_t^{t+\tau} s(z)dz \le K_\tau, \quad \int_t^{t+\tau} s^{(i)}(z)dz \le K_\tau^{(i)}, \quad \sum_{i=0}^{n}\int_t^{t+\tau} s^{(i)}(z)dz \le M_\tau$$

Note that the fulfillment for $\tau > 0$ of one of the two conditions

$$\int_t^{t+\tau} s^{(i)}(z)dz > K_\tau^{(i)}, \quad \sum_{i=0}^{n}\int_t^{t+\tau} s^{(i)}(z)dz > M_\tau$$

leads to the destruction (catastrophe) of the system.

Example. Let there be some socio-economic environment that renews its resources with a renewal rate $\alpha(\tau, t, x)(0 < t < T,\ 0 < X < 1,\ 0 < \tau < T)$. This coefficient depends, in general, on the power of the environment (its resource intensity, resource availability). Consider a simple hypothesis: $\alpha(\tau, t, x) = \alpha_0 +$

$\alpha_1 x$, and the more resources, the greater the rate of their renewal. You can write down a continuous evolutionary model (a is the coefficient of natural increase in resources, b is their loss):

$$x'(\tau) = -\left[a(\tau) + \int_0^\tau b(\tau, s)x(s)ds\right]x(\tau),$$

$$x(0) = \int_0^T [\alpha_0(\tau) + \alpha_1(\tau)x(\tau)]\, x(\tau)d\tau\,.$$

Let $\alpha(\tau, x) = \alpha_0(\tau) + \alpha_1(\tau)x(\tau) > 0$.. Then

$$x(0) = \int_0^T \alpha(\tau, x)\, x(\tau)d\tau\,.$$

The problem always has a solution x ≡ 0. Then the evolutionary potential of the system can be defined as the value:

$$\lambda = \int_0^T \alpha(\tau)e^{-\int_0^\tau a(s)ds}d\tau.$$

The higher the tempo of α, the higher the λ, the lower the α, the lower the λ. No matter how good the state of the resources at the initial moment, they will invariably be depleted if the potential of the system is less than 1.

Example. Let u_{max} be the maximum level of syntax errors in the program P, u(t) – their remaining amount by the time t. Based on the simplest evolutionary model $du/dt + \lambda u_{max} = 0,\ u(i_0) = u_0,$ we can conclude that the error rate decreases as $\lambda(c - t_0) \neq -1\ (t_0 < c < T)$ according to the law: $u(t) = u_0\,(1 + \lambda(c - t))/(1 + \lambda(c - t_0))$. If you additionally set $u(t_*) = u_*$, (u_{max} is an unknown quantity, $t_* \neq t_0$), then the law of change in the level of errors can be found unambiguously, since: $c = (u_* t_0 - u_0 t_*)/(\lambda u_* - \lambda u_0) - 1/\lambda$.

Note that if ds/dt is the total change in the entropy of the system when acting on the system, ds_1/dt is the change in entropy due to irreversible changes in the structure, flows within the system (considered as an open system), ds_2/dt is the change in entropy due to efforts to improve situation (for example, economic, environmental, social), then the equation of I. Prigogine is valid:

$$ds/dt = ds_1/dt + ds_2/dt.$$

In evolutionary modeling of socio-economic systems, it is useful to use both classical mathematical models and non-classical ones, in particular, taking into account the spatial structure of the system (for example, cellular automata and fractals), the structure and hierarchy of subsystems (for example, graphs and data structures), experience and intuition (for example, heuristic, expert procedures).

Example. Let some ecological system Ω be given, in which there are points of pollution (emissions of pollutants) x_i, i = 1, 2, …, n. Each pollutant x_i sequentially pollutes the ecosystem over a period of time $(t_{i-1}; t_i]$, $t_i = t_i – t_{i-1}$. Each pollutant can affect the activity of another pollutant (eg, reduce, neutralize or enhance by the known effect of summation of the effects of pollutants). The strength (measure) of such influence can be defined through r_{ij}, $R = \{r_{ij}: i = 1, 2, …, n – 1; j = 2, 3, …, n\}$.

The structure is given by the graph: vertices – pollutants, edges – measurees.

Let us find a substitution that minimizes a functional of the form:

$$F = \sum_{i=1}^{n-1} \sum\nolimits_{j=2}^{n} r_{ij} = min,$$

where F is the total pollution of the system with a given structure S.

The faster (slower) the accounting of pollution at point x_i, the faster (slower) socio-economic measures to neutralize it (increase the impact) are feasible. The less pollutants there are before pollutant x_i, the less pollution will be.

As a measure of r_{ij}, a measure can be taken that takes into account both the time of onset of exposure to pollutants (preceding this x_j) and the number and intensity of these pollutants:

$$s_{ij} = v_{ij}\left[h_j \sum_{k=i+1}^{j-1} \tau_k + (1 + h_j)(j - i)\right]$$

where v_{ij} is the weight factor that determines the degree of influence of the pollutant x_i on the pollutant x_j (summation effect), h_j is the weight factor that takes into account the specific intensity of the effect of the pollutant x_j or the τ_i interval during which the intensity (concentration) of the pollutant decreases. Weighting factors are established by experts or experimentally.

The principle of evolutionary modeling presupposes the necessity and efficiency of using the methods and technologies of artificial intelligence, in particular, expert systems.

The main difficulty in the construction and use of evolutionary models: in Nature and Cognition, in which these models and goals exist explicitly or implicitly, the results of the functioning of the system and the achievement of the goal are often traceable only after a long period of time, although in Society and Economy Man seeks to obtain results in compliance with the goal clearly and quickly, with a minimum cost of resources.

2 Genetic algorithms

Genetic algorithms are an adequate means of implementing evolutionary modeling procedures.

The idea of genetic algorithms is "spied" from systems of living nature, from systems whose evolution is developing in complex systems rather quickly.

A genetic algorithm is an algorithm based on imitating the genetic procedures for the development of a population in accordance with the principles of evolutionary dynamics given above. It is often used to solve problems of optimization (multi-criteria), search, management.

These algorithms are adaptive, develop solutions, develop themselves. A feature of these algorithms is their successful use in solving NP-hard problems (problems for which it is impossible to construct an algorithm with polynomially increasing algorithmic complexity).

Example. Consider the problem of unconstrained integer optimization (placement): find the maximum f(i), i is a set of n zeros and ones, for example, for n = 5, i = (1, 0, 0, 1, 0). This is a very difficult combinatorial problem for conventional, "non-genetic" algorithms. The genetic algorithm can be constructed by the following generalized procedure:

Listing 12.1

1 генерируем начальную популяцию (набор допустимых решений задачи) – $I_0 = (i_1, i_2, \ldots, i_n)$, $i_j \in \{0, 1\}$ и определяем некоторый критерий достижения "хорошего" решения, критерий остановки α, процедуру СЕЛЕКЦИЯ, процедуру СКРЕЩИВАНИЕ, процедуру МУТАЦИЯ и процедуру обновления популяции ОБНОВИТЬ;
2 $k := 0,\ \ f_0 := max\{f(i),\ i\epsilon I_0\}$
3 начало цикла пока не(α)
    1 с помощью вероятностного оператора (селекции) выбираем два допустимых решения (родителей) i_1, i_2 из выбранной популяции (вызов процедуры СЕЛЕКЦИЯ);
    2 по этим родителям строим новое решение (вызов процедуры СКРЕЩИВАНИЕ) и получаем новое решение i;
    3 модифицируем это решение (вызов процедуры МУТАЦИЯ);
    4 если $f_0 < f(i)$ то f0 := f(i);
    5 обновляем популяцию (вызов процедуры ОБНОВИТЬ);
    6 k := k + 1
4 конец цикла

These procedures are determined using similar procedures of living nature (at the level of knowledge about them that we have). For example, the SELECTION procedure can select from random elements of the population the element with the highest value of f(i). The procedure CROSSING (crossover) can construct a vector i from vectors i_1, i_2, assigning with a probability 0.5 the corresponding coordinate of each of these parent vectors. This is the simplest procedure. More complex procedures are also used that implement more complete analogs of genetic mechanisms. The MUTATION procedure can also be simple or complex. For example, a simple procedure with a given probability for each vector reverses its coordinates (0 by 1, and vice versa). The UPDATE procedure is to update all elements of the population in accordance with the specified procedures.

Example. The bank's operation can be simulated based on genetic algorithms. With their help, you can choose the optimal bank interest (deposits, loans) of a bank in a competitive environment in order to attract more customers (funds). The bank that will be able to attract more deposits, clients and funds, and develop a more attractive strategy of behavior (evolution), will survive under the conditions of natural selection. The branches of such a bank (genes) will better adapt and strengthen in the economic niche, and, possibly, will increase with each new generation. Each bank branch (individual of the population) can be estimated by the measure of its fitness. Such measures may be based on various criteria, for example, an analogue of economic potential – a bank's reliability rating or the ratio of borrowed and own funds of a bank. Such an assessment is equivalent to assessing how efficient an organism is in competing for resources, i.e. its survival, biological potential. In this case, individuals (branches) can lead to the appearance of offspring (new banks obtained as a result of merger or disintegration), combining certain (economic) characteristics of the parents. For example, if one bank had a good lending policy and the other had an effective investment policy, then the new bank could acquire both. The least adapted individuals (branches) may completely disappear as a result of evolution. Thus, the genetic procedure for the reproduction of new banks (new generation), more adapted and able to survive in the evolution of the banking system, is being worked out. Over time, this policy permeates the entire banking "population", ensuring the achievement of the goal - the emergence of an efficient, reliable and stable banking system. Here is the corresponding genetic algorithm (enlarged and simplified):

Listing 12.2

```
алг ГЕНЕТИЧЕСКИЙ_АЛГОРИТМ_БАНКОВСКОЙ_СИСТЕМЫ
  ввод Начальная структура банка (начальная популяция);
  СТРУКТУРА | процедура оценки структуры по приспособлению
  Стоп:=0   | флаг для завершения эволюционного процесса
  нц пока (Стоп=0)
        СЕЛЕКЦИЯ | процедура генетического отбора нового поколения
        нц пока (МЕРА) | цикл воспроизводства с критерием МЕРА
                  | мерой эффективности банковской системы
              РОДИТЕЛИ  | процедура выбора двух структур (филиалов)
                    | объединяемых (скрещиваемых) на новом шаге
              ОБЪЕДИНЕНИЕ | процедура образования (объединения)
                      | нового банка (филиала)
              ОЦЕНКА | процедура оценки устойчивости нового банка,
                    | образования (рейтинга, устойчивости)
              ВКЛЮЧЕНИЕ | процедура включения (не включения) в новое
                    | поколение (в банковскую систему)
        кц
```

```
        МУТАЦИЯ | процедура эволюции (мутации) нового поколения
        если (ПРОЦЕСС) | проверка функционала завершаемости эволюции
            то Стоп:=1
  кц
кон.
```

We do not specify the structure of the procedures SELECTION, MEASURE, PARENTS, UNION, EVALUATION, INCLUSION, MUTATION, PROCESS, although even at the intuitive level it is clear that in this algorithm they play a decisive role for the evolutionary process. The correct (effective) choice of the structure, as well as the presentation (description) of this structure, is no less important. It is often chosen by analogy with the structure of chromosomes, for example, in the form of bit strings. Each string (chromosome) is a concatenation of a number of substrings (gene combination). Genes are located at different positions on the line (chromosome loci). They can take on some values (alleles), for example, for bit representation - 0 and 1. The data structure in the genetic algorithm (genotype) reflects the genetic model of an individual. Environment, environment is determined by a vector in the parameter space and corresponds to the term "phenotype". The measure of quality (MERA procedure) of a structure is often determined by an objective function (fitness). For each new generation, the genetic algorithm carries out selection proportionally to fitness (SELECTION procedure), modification (Parents, UNION, INCLUSION procedures) and mutation (MUTATION procedure). For example, in the SELECTION procedure, each structure is associated with the ratio of its fitness to the total fitness of the population, and then selection (with replacement) of all individuals takes place for further genetic processing in accordance with this value. The size of the selected combination can be taken proportional to the fitness, and therefore individuals (clusters) with higher fitness are more likely to be selected than individuals with low fitness. After selection, the selected individuals undergo crossover (recombination), i.e. split into pairs. A crossover can be applied to each pair. Unchanged individuals move to the stage of mutation. If a crossover occurs, the resulting offspring replace their parents and proceed to mutate.

Although genetic algorithms can be used to solve problems that, apparently, cannot be solved by other methods, they do not guarantee finding the optimal solution (at least, in an acceptable time; polynomial estimates are often inapplicable here). Criteria like "good enough and fast enough" are more appropriate here. The main advantage is different: they allow solving complex problems for which stable and acceptable methods have not yet been developed, especially at the stage of formalization and structuring of the system, in cognitive systems. Genetic algorithms are effective in combination with other classical algorithms, heuristic procedures, as well as in cases where there is some additional information about the set of decisions, which allows you to adjust the parameters of the model, adjust the selection and evolution criteria.

Conclusion

The basic concepts and principles of evolutionary modeling of systems are considered, as well as genetic algorithms – an adequate apparatus for its implementation.

Control questions

See the manual on the organization of students' independent work.