MÜNCHENER STOCHASTIK-TAGE 1998 Sekt.-Nr.:14

Stochastische Optimierung und Neuronale Netze

Sektionsleiter: K. Marti, München; M. H. Breitner, Clausthal-Z.

Breitner, Michael H.
Institut für Mathematik / TU Clausthal

Efficient Neural Network Training and Neurocomputing

When real-life applications are treated, neural networks are typically computationally expensive because

- the dimension of the input and/or output vector is large;
- the pattern set is huge and/or the pattern set must be split appropriately into subsets which are used to train different neural subnetworks;
- no convergence to the global minimum of the error function $\varepsilon_t(W)$ can be ensured (training with very many random weight initializations necessary).

Acceptable computing times can be achieved on the one hand by specialized neuro-computers which support the parallel architecture of neural networks in hardware. We focus on the SYNAPSE2[1] neuro-computer based on the patented MA16 neuro-signal processor. The SYNAPSE2 has a peak performance of 1.28 billion multiplications (16 bit fixed-point)/additions (48 bit fixed-point) per second plus 40 million look-up-table calls (16 bit fixed-point) of predefineable functions. For multiplication and addition of 16x4 (sub-)matrices and adjacent evaluation of only few functions the compute power of the SYNAPSE2 can be about two to three orders of magnitude above conventional workstations. Here the emphasis is on multi-layer perceptrons found in many successful applications, both research and commercial. Appropriate design of various network architectures and implementation of learning algorithms with the C++ class library SynUse•Base are discussed. Only basic training algorithms (backpropagation and momentum learning) can be implemented easily for the SYNAPSE2. On the other hand supervised learning is possible with all methods for constrained nonlinear least-squares problems. Advanced methods are sequential quadratic programming (SQP) methods and generalized Gauß-Newton (GGN) methods which can exploit the special structure of the Hessian of $\varepsilon_t(W)$. SQP and GGN methods approximate the Hessian by finite-differences and can deal with box constraints on the weights, linear constraints and smooth nonlinear constraints. Advantages of these methods are:

- The step length for training is optimized permanently. In comparison to common training methods with fixed step length the number of learning steps is reduced significantly (factor 10 – 1000).
- Only first derivatives of $\varepsilon_t(W)$ are required which can be computed very fast. The first derivatives can be deduced easily by automatic differentiation (backpropagation).
- Maximum and minimum of each weight can be set easily (box constraints).
- The total curvature of the neural network can be constrained (prevention of neural network oscillations).
- SQP and GGN methods often can automatically overcome training problems of three-layer perceptrons, e. g., flat spots or steep canyons of $\varepsilon_t(W)$.

Note that usually the training stops before a local minimum of $\varepsilon_t(W)$ is reached for all training methods. For large training sets often thousands, ten thousands or more local minima exist even

[1] = Synthesis of neural algorithms on a parallel systolic engine.

for three-layer perceptrons with only few input and hidden neurons. The more neural networks with different random weight initializations are trained, the more likely it is to train a successful perceptron, i. e. to find a good local minimum of $\varepsilon_t(W)$.

Cai, Jianbo
Fachgebiet Statik/Baumechanik, FB 10
Universität GH Essen, 45117 Essen

Stochastische Suchmethoden in der Strukturoptimierung

In diesem Beitrag werden die Grundlagen zweier stochastischer Suchmethoden der Strukturoptimierung, Evolutionsstrategien und genetische Algorithmen, dargestellt. Diese Methoden bilden die biologische Evolution nach und arbeiten mit Zufallsoperatoren wie *Mutation*, *Rekombination* und *Selektion*. Der einfache Suchmechanismus erlaubt eine sehr breite Anwendung dieser Methoden im Ingenieurbereich. Anstelle eines einzelnen Bemessungspunktes behandeln diese Methoden eine Population von Bemessungspunkten im Variablenraum. Diese Eigenschaft ermöglicht die Parallelisierung dieser Methoden. Parallele Strategien zur Lösung von gemischt-diskreter Optimierungsprobleme werden vorgestellt. Anwendungsbeispiele im Bauwesen werden präsentiert.

Dippon , Jürgen
Universität Stuttgart

Beschleunigte stochastische Optimierung mit zufälligen Richtungen

Die Minimalstelle ϑ einer unbekannten Regressionsfunktion $f : \mathbf{R}^d \to \mathbf{R}$ kann rekursiv durch

$$X_{n+1} = X_n - a_n Y_n$$

geschätzt werden, wobei Y_n eine Schätzung des Gradienten von $f(X_n)$ darstellt und (a_n) eine geeignete Nullfolge ist. Für den *finite difference* Gradientenschätzer von Kiefer und Wolfowitz (1952), der $2d$ Beobachtungen benötigt, kann bei dreifacher Differenzierbarkeit von f an der Stelle ϑ die asymptotische Normalität von $n^{1/3}(X_n - \vartheta)$ gezeigt werden. Letzteres gilt auch für den *random-direction* Schätzer von Kushner und Clark (1978) und den *simultaneous-perturbation* Schätzer von Spall (1992), welche jedoch nur jeweils zwei Beobachtungen je Iterationsschritt benötigen. Die Grenzverteilungen sind i.a. bei allen drei Verfahren verschieden und auch verzerrt. Ein Effizienzvergleich ist deshalb schwierig.

In dem Vortrag stellen wir Varianten dieser Verfahren vor, deren Grenzverteilung unverzerrt ist, falls statt zwei (bzw. $2d$) Beobachtungen vier (bzw. $4d$) je Rekursionsschritt zur Verfügung stehen. Besitzt f an der Stelle ϑ Ableitungen der Ordnung $p \geq 3$, kann unter Verwendung von $2\lceil p/2 \rceil$ (bzw. $2d\lceil p/2 \rceil$) Beobachtungen die Konvergenzordnung $n^{-\frac{1}{2}(1-\frac{1}{p})}$ erreicht werden.

Eberhard, Peter
Institute B of Mechanics, University of Stuttgart, Germany

Is Simulated Annealing Useful for Complex Multibody Systems?

During the last decades, the multibody system approach for the investigation of mechanical systems with large motions evolved to a well accepted and widely used method. Highly increased computational power and advances in the theoretical understanding allowed to start changing the focus from a pure analysis to system synthesis, i.e. after modeling and simulating the system behavior we try to tune the system parameters in order to optimally fulfil requirements and system specifications.

The dynamics of a stochastically excited multibody system can be described by first order differential equations

$$\begin{aligned} \dot{\mathbf{y}} &= \mathbf{v}(t,\mathbf{y},\mathbf{z},\mathbf{p}), & \mathbf{y}(t_0) &= \mathbf{y}_0(\mathbf{p}), \\ \mathbf{M}(t,\mathbf{y},\mathbf{p})\dot{\mathbf{z}} + \mathbf{k}(t,\mathbf{y},\mathbf{z},\mathbf{p}) &= \mathbf{q}(t,\mathbf{y},\mathbf{z},\mathbf{p},\boldsymbol{\xi}), & \mathbf{z}(t_0) &= \mathbf{z}_0(\mathbf{p}), \\ \dot{\boldsymbol{\xi}}(t) &= \mathbf{C}\boldsymbol{\xi}(t) + \mathbf{D}\mathbf{w}(t), & \boldsymbol{\xi}(t_0) &= \boldsymbol{\xi}_0 \end{aligned}$$

where $\mathbf{y}$ is a vector of generalized coordinates and $\mathbf{z}$ is a vector of generalized velocities. The colored noice $\boldsymbol{\xi}$ is
generated from white noise $\mathbf{w}$ by a linear formfilter. The values of the design parameters $\mathbf{p}$ have to be chosen appropriately by an optimization procedure in order to minimize a set of performance criteria.

Numerical investigations in the field of stochastic excited vehicles show that deterministic scalar optimization algorithms like SQP-methods tend to stop in local minima far away from the global optimum. Stochastic optimization algorithms like simulated annealing methods may offer a way to perform global optimization although they require much more iterations. Since not only a single point of the EP-optimal set is generated, they also give a clearer picture what can be reached in a multicriteria optimization problem.

After a short overview on the analysis and optimization concepts, we will present a hybrid stochastic/deterministic optimization approach with application to stochastic multibody systems. Some comparisons of stochastic, deterministic and hybrid methods applied to a complex nonlinear spatial vehicle under stochastic excitation will be shown and evaluated.

Egorov , Igor
Air Force Engineering Academy, Moskau, Russia

Indirect Optimizaion Method on the Basis of Self-Organization. Destination, Main Algorithm and Numerical Testing.

The aim of this paper is to discuss a new numerical technique aimed at solving a wide variety of complex technical system optimization problems. The main advantage of this technique consists of maximum economies of a number of direct calls to a mathematical simulator of the system under consideration during search of the solution. In the paper a brief description of the basic algorithm of the proposed optimization method and examples of its numerical testing are given. The method's effectiveness has been checked on 13 test functions, which have been selected according to imitation of various types of goal functions complexity.The main concept of proposed method is to store and use with maximal effectiveness the information about goal function, which is being accumulated in the process of problem solving. The central place of our method is the presentation of each solution search iteration as two stages: to construct the functions approximating optimization criterion and functional constraints in some chosen region and to find this approximating function extreme with constraints taken into account. To realize such strategy the unique procedures have been developed to approximate functions of many variables (using the inductive self-organization method of academician A.G. Ivakhnenko, which has been modified for optimization problems). The effectiveness of these procedures has been proved by solving of considerable number of testing and practical optimization problems.The developed method permits us to solve a number of optimization problems for the real gas-turbine engines and their units, which have been constructed in the Russian scientific-production organizations and also in several foreign engine-building firms. The method is invariant to the objects under study and can be adopted for the use in different spheres of science and engineering.

Ermoliev , Yuri
IIASA, International Institute for Applied Systems Analysis, Austria

Insurability of Catastrophic Risks: The Role of Stochastic Optimization Models

The talk is concerned with the development of decision making tools for managing catastrophic (low probability - high consequences) risks. Insurance, one of the main mechanisms against risks, is currently concerned with the possibility of catastrophic claims even higher than already experienced. The principal problem in insuring catastrophic risks is insufficient historical data for predicting events at a particular location, although rich data may exist on their occurrence and magnitude on an aggregated (say regional) level. Potential damages in a particular location may be unlike anything that has been experienced in the past. Catastrophes produce highly correlated damages and claims, which depend on the region of occurrence, coverages at different locations, mitigation measures, reinsurance agreements and so on. Joint probability distributions of these claims depicting their complex spatial and temporal interactions and effects of decision variables are analytically intractable. The lack of historical data and the complex spatial and dynamic interdependencies make it dangerous to use purely adaptive "trial-and-error" approaches. In such cases the role of models becomes crucially important. Spatial dynamic stochastic optimization models have been developed enabling one to incorporate the simulation of realistic and geographically explicit catastrophic events, property values and insurance coverages in different locations of the region. Combined with stochastic optimization techniques they aid decision making on the spatial diversification of contracts, premiums, reinsurance requirements, effects of mitigation measures, regulations and the use of other financial mechanisms. The models involve stopping times, extreme events and nonsmooth dependent risk accumulation processes. In particular they allow one to deal with catastrophes caused by the clustering in time of such events as rains and droughts.

Franke, Jürgen; Neumann, Michael,
Universität Kaiserslautern
Humboldt-Universität zu Berlin

Bootstrap-Verfahren für neuronale Netze

Zur Motivation wird kurz vorgestellt, wie Bootstrap-Ansätze bei der Modellierung von Finanzdaten mit neuronalen Netzen eingesetzt werden. Anschließend zeigen wir, dä verschiedene Bootstrap-Verfahren (residuenbasiertes, paarweises und Wild Bootstrap) trotz der im Normalfall vorliegenden Misspezifikation des Modells für neuronale Netze angewandt werden können, und diskutieren einige spezifische, technische Probleme, die bei der praktischen Verwendung auftreten können.

Frauendorfer , Karl
Universität St. Gallen, Institut für Unternehmensforschung (IfU-HSG), Stochastische Optimierung

Stochastic Multistage Programming: Experiences in Finance

The complexity of the interaction between time and uncertainty made finance models to one of the most important applications of probability and optimization theory. Stochastic multistage programming combines these two fields with the intention to design sophisticated, analytical tools for analyzing interaction effects between decision making and uncertainty.
In this work we report on the experiences we have achieved in the last five years with applying stochastic multistage programming to investment and funding problems, to cash management and to asset allocation. Modelling, solution techniques and the role of convexity will be revealed in detail. Modelling covers how interest rate models are incorporated adequately so that the correlation between interest rates and uncertain cash flows are exploited, and how various sources of

uncertainty inherent to non-maturing accounts like prepayment or withdrawal risk may be taken into account. Within the problem of asset allocation the classical mean-variance criterion has been extended to a multiperiod setting which allows for incorporating the dynamcis of variances and covariances and rebalancing activities. Barycentric approximation is used as solution technique for solving the stochastic multistage programs. This approach discretizes the underlying stochastic processes with respect to time and state. In case that the problems are convex the accuracy of the distribution may be assessed and the inaccuracy of the approximation localized. This contributes to derive refinement activities that enhance an appealing convergence behaviour.
Latest numerical results will be presented that have been achieved from various case studies with the banking and insurance industry.

Gallegos, Maria Teresa
Fakultät für Mathematik und Informatik, Universität Passau

Konvergenz des Verstärkungslernens für eine allgemeine Lösungsmenge

Verstärkungslernen (reinforcement learning) ist das Lernen über Beispiele. Einige seiner Konzepte wurden ursprünglich von mathematisch–orientierten Psychologen und Biologen entwickelt, um das Lernverhalten von Individuen sowohl vom psychologischen als auch vom biologischen Standpunkt aus zu beschreiben.
Verstärkungslernen bedeutet "die Wahrscheinlichkeit einer Reaktion zu erhöhen" (increasing the probability of a response). Das Lernsystem, ein Automat, bietet der Umwelt eine Lösung aus einer gegebenen endlichen Menge von Lösungsmöglichkeiten an. Die Umwelt beurteilt die Lösung, z.B. nur durch Mitteilung "gut" oder "schlecht". Das System paßt dann seinen aktuellen Zustand nach der Beurteilung aufgrund einer Verstärkungsregel an und versucht so, einen optimalen Zustand zu erreichen. Die Zustandsfolgen dieser Algorithmen werden durch homogene Markov–Ketten beschrieben. Das Hauptproblem des Lernens besteht darin, Bedingungen über den Algorithmus anzugeben, so daß die zugehörige Markov–Kette das gewünschte asymptotische Lernverhalten aufweist. Eines der Hauptziele des Vortrages ist es, eine Verallgemeinerung eines von Oommen angegebenen Algorithmus für beliebig–viele reine Lösungen anzugeben und mathematisch zu analysieren. Damit wird auch eine von Oommen und Christensen geäußerte Vermutung bestätigt. Der Beweis beruht auf Vergleichssätzen für Verteilungen und Abschätzungen von Gleichgewichten.

Grill , Heiner
Lehrstuhl für Ingenieurinformatik im Bauwesen, Graduiertenkolleg Computational Structural Dynamics, Ruhr-Universität Bochum

Diskret-kontinuierliche Strukturoptimierung mit verteilten Evolutionsstrategien

Dieser Beitrag befaßt sich mit dem Einsatz von Evolutionsstrategien zum gewichtsminimalen Entwurf von Tragstrukturen unter Berücksichtigung sowohl kontinuierlicher als auch diskreter Entwurfsvariablen.
Evolutionsstrategien sind, ähnlich wie auch Genetische Algorithmen, direkte Optimierungsverfahren, deren Funktionsweise sich am Vorbild der natürlichen Evolution orientiert. Diese beruht auf dem Einsatz der teilweise zufallsbasierten Evolutionsoperatoren Rekombination, Mutation und Selektion. Anstelle der sequentiellen Betrachtung einzelner Entwurfspunkte, wie bei ableitungsorientierten Verfahren meist üblich, wird hier das Lernvermögen einer Population simultan und in Konkurrenz zueinander existierender Entwurfsvarianten, sog. Individuen, im Entwurfsraum genutzt. Im Zusammenspiel der Evolutionsoperatoren bewirkt der Rekombinationsoperator den Austausch von Entwurfsinformationen zwischen bestehenden Entwurfsvarianten, der Mutationsoperator dagegen erzeugt neue Entwurfsinformationen durch normalverteilte Zufallsänderungen vorliegender Individuen. Durch den Selektionsoperator werden die qualitativ hochwertigsten Individuen der

aktuell betrachteten Population als Ausgangsbasis zur Bildung der Population im nachfolgenden Evolutionsschritt ausgewählt.
Evolutionsstrategien stellen insgesamt eine ist günstige Möglichkeit zur Entwicklung robuster und hochgradig problemneutraler Optimierungssysteme dar. Als entscheidender Nachteil, insbesondere im Bereich der Strukturoptimierung, wurde bislang häufig der hohe Rechenaufwand angesehen. Die Eigenschaft, gleichzeitig eine Vielfalt unabhängiger Individuen in einer Population zu betrachten, ermöglicht allerdings eine effiziente Realisierung in zunehmend verfügbaren, verteilten Hardware-Umgebungen. Hierfür kommen sowohl Gruppen vernetzter Arbeitsplatzrechner als auch dedizierte, evtl. hochgradig parallele Multiprozessorsystemen in Betracht. Es wird gezeigt, wie durch geeignete Maßnahmen zur Lastverteilung eine effiziente Nutzung unterschiedlicher Systemumgebungen sichergestellt und eine erhebliche Reduktion der Rechenzeiten erzielt werden kann.

Gröwe-Kuska , Nicole
Institut für Mathematik, Humboldt-Universität Berlin

Estimated stochastic programs with chance constraints

Stochastic programs with chance constraints are specifically structured programs that are derived from optimization problems involving random parameters. In most applications of stochastic programming we have to estimate the unknown probability distribution of the random parameters from empirical data.
We propose a nonparametric estimation procedure for chance-constrained stochastic programs where the random parameters appear on the right-hand side of linear constraints for the decision variable. The assumed independence of the components of the random right-hand side data results in stochastic programs with a separability structure in the constraints. Motivated by the relationship between logarithmic concave measures and increasing hazard rate distributions we estimate the unknown distribution function of the random data via isotonic regression estimates of increasing hazard rates. By our choice of the estimates the feasible region of the estimated stochastic programs can be described by linear equations.
We study asymptotic properties of optimal values and optimal solutions of the estimated stochastic programs if the size of the random samples tends to infinity. Asymptotic confidence intervals for optimal values and asymptotic weak confidence sets for optimal solutions are derived from large deviation type results.
Finally, we discuss the numerical treatment of the estimated chance-constrained programs and report on test runs.

Jendo, Stefan; Niczyj, Janusz
Polish Academy of Sciences, Warsaw, Poland;
Technical University of Szczecin, Szczecin, Poland

Reliability assessment of existing structures using fuzzy logic

Most structures are composed of many members. The reliability of a structural system will be a function of the reliability of its members. Any real structural systems usually requires a combination of both serial and parallel description of its subsystems. The paper deals with reliability of a plane truss structures. The linguistic values of elements reliability are obtained by experiments and expert opinions. The final computational steps is to determine a single linguistic estimate of the state of the construction. The difficult problem of damage assessment of structures is subdivided into smaller problems for each element. From the static scheme of the structure we derive a system of elements according to the theory of reliability. This model gave us the information about the influence of elements damage on the state of the whole structure. We assume eight linguistic terms for damage of an element: none, slight, small, moderate, considerable, severe, very severe and destructive. In

the paper, we show how the system may be described, that is, the way in which the state of the system depends on the state of its elements. Finally the study of reliability functions of complex systems of truss structures is presented, leading to a fuzzy reliability of structures. In conclusion, the major problem addressed in this paper is the presentation of fuzzy system failure probability by fuzzy linguistic variables and by a fuzzy number over an interval of confidence. The use of interval arithmetic and alfa-cuts gives considerable advantage for evaluation of fuzzy systems. In this study we have shown that by descibing uncertainty of damage state of elements with fuzzy sets one may estimate the reliability of existing construction in case of incomplete or uncertain knowledge of its parameters. The proposed method may be used in expert systems.

Kan , Yuri
Moscow Aviation Institute

THE CONVEXITY OF THE QUANTILE FUNCTION

The quantile function plays an important role in Stochastic Programming. Its convexity is very useful when we deal with the optimization models including probabilistic constraints.
The quantile of level $\alpha \in (0,1)$ for the random variable η is denoted by

$$(\eta)_\alpha = \min\{z|\ P(\eta \le z) \ge \alpha\}.$$

Let $f(u,\xi)$ be a function which is Borel-measurable with respect to the random vector ξ. Then the quantile function is defined as follows:

$$f_\alpha(u) = (f(u,\xi))_\alpha.$$

It is known [1] that $f_\alpha(u)$ is convex in u if $f(u,\xi)$ is jointly quasi-convex in u and ξ. The latter condition seems rather hard, since in applications we often find loss functions which are quasi-convex both in u and ξ but not jointly quasi-convex in u and ξ. Nevertheless the quantile function is expected to be convex. For example, in [2] it is factically proved that $(u^T\xi)_\alpha$ is convex if ξ has a multi-variate normal distribution.
We introduce into consideration the α-kernel

$$K_\alpha = \bigcap_a \left\{x|\ a^T x \le (a^T\xi)_\alpha\right\}.$$

K_α is called to be *regular* if it is non-empty and every closed half-space H containing the α-kernel is confidence, i.e. $P(H) \ge \alpha$.

Theorem *Let $\alpha \in (0,1)$ be a confidence probability and η and ψ be random variables such that for the random vector (η,ψ) the α-kernel is regular. Then*

$$(\eta+\psi)_\alpha \le (\eta)_\alpha + (\psi)_\alpha.$$

If for all feasible u_1 and u_2 this theorem is applicable for $\eta = f(u_1,\xi)$ and $\psi = f(u_2,\xi)$ it immediately follows that the quantile function $f_\alpha(u)$ is convex in u.
Sufficient conditions for the regularity of the α-kernel and the applicability of the theorem are discussed.

Literatur

[1] Kibzun A.I. and Yu.S.Kan *Stochastic programming problems with probability and quantile functions.* John Wiley & Sons, 1996.

[2] Kall P. *Stochastic linear programming.* Springer-Verlag, 1976.

Lenzen , Armin
Lehrstuhl für Ingenieurinformatik im Bauwesen, Ruhr - Universität - Bochum

Stochastische Strukturanalyse zur lebensdauerorientierten Optimierung

Der Entwurf von Konstruktionen des Ingenieurbaus kann als ein Optimierungsprozeß verstanden werden, bei dem eine Vielzahl von planungsrelevanten Größen zu erfüllen sind bzw. Eingang finden, die stochastischen Charakter haben. Hierbei sind insbesondere Belastungen der Bauwerke aus der Natur (z. B. Wind, Meereswellen, Erdbeben) zu nennen, die sich nicht in einfacher deterministischer Form behandeln lassen. Aber auch Parameter die in das mechanische Strukturmodell einfließen, wie z.B. Stoffgrößen, haben streuenden Charakter.
Die genannten Belastungen können als Zufallsprozesse modelliert werden, und werden so einer mathematischen Beschreibung zugänglich gemacht. Der gemessene mehrfachkorrelierte Lastprozess wird über eine numerische Methode (Realisierung mit Singulärwertzerlegung) in ein mathematisches Parametermodell überfuehrt. Auf Basis der Systemtheorie wird der Belastungsprozeß mit der Finite Elemente Methode, zur Beschreibung der Struktur, zu einer stochastischen Strukturanalyse gekoppelt.
Der farbige Prozeß der Strukturerregung wird auf weißes Rauschen zurückgeführt und damit nach Reihenschaltung von Lastfilter und Strukturfilter der bekannten Lösung einer Ito - DGL zugänglich. Das numerische Verfahren, das auch als Kovarianzanalyse bekannt ist, ist der Time-History und der Spektralmethode in Rechenzeitbedarf überlegen, da eine algebraische Gleichung (Ljapunov) zur Loesung der stochastischen DGL besteht.
Die Ergebnisse der Analyse erlauben dem Ingenieur eine Beurteilung und Auslegung unter wahrscheinlichkeitstheoretischen Gesichtspunkten.
Diese rechenzeitgünstige Methode soll in Zukunft als Kern eines stochastischen Optimierungsmodells zur lebensdauerorientierten Auslegung Verwendung finden.
Vortragsinhalt ist die Erläuterung der Methode und eine Anwendung an einem Beispiel aus der Praxis (Stahlrahmen mit Windbelastung).

Marti, Kurt; Stoeckl , G.
UniBw München, Fak. LRT, 85577 Neubiberg/München

Entwurfsoptimierung von Tragwerken mit Hilfe der zweistufigen Stochastischen Linearen Programmierung

Strukturoptimierung von Tragwerken mit stochastischen Modellparametern (Materialdaten, Last, usw.) erfordert, daß für das ursprüngliche stochastische Strukturoptimierungsproblem zunächst ein geeignetes deterministisches Ersatzproblem formuliert wird. Ausgehend von den linearen Ungleichungen (Festigkeitsbedingungen) für die inneren Kräfte in den Elementen des Tragwerkes und deren Gleichgewichtsbedingungen kann das Problem als 'Two-Stage Stochastic Linear Program with Complete Fixed Recourse' beschrieben werden. Dabei ergeben sich die W-Matrix und die Kostenfaktoren q für die zweite Stufe durch Einbeziehung unerwünschter Veränderungen der Elemente des Tragwerks (verursacht durch die inneren Kräfte) oder durch Betrachten bestimmter Festigkeits- oder Volumenreserven. Der optimale Entwurf (optimal design) kann dann für beliebige diskrete Wahrscheinlichkeitsverteilungen der stochastischen Modellparameter bestimmt werden durch Lösen eines besonders strukturierten, evtl. sehr grossen linearen Programms (LP). Nachdem im 1. Teil verschiedene deterministische Ersatzprobleme für Entwurfsoptimierungsprobleme unter Unsicherheit vorgestellt werden, erfolgt im 2. Teil die Präsentation numerischer Ergebnisse für einfache Tragwerke (trusses) und verschiedene Ersatzprobleme. Ferner wird die Abhängigkeit der Optimallösung von der zugrundeliegenden diskreten Verteilung der Zufallsgrößen diskutiert.

Mayer, János
Institute for Operations Research, University of Zurich

On the role of bounds in stochastic programming algorithms

This is a joint lecture with Prof. Dr. Peter Kall, IOR University of Zurich.
We consider the following types of stochastic linear programming (SLP) problems: jointly chance constrained problems and SLP problems with recourse. These problems are numerically hard to solve. The numerical difficulties have their roots in multivariate integrals in the problem formulations. For chance constrained problems multivariate probability distribution functions and their gradients are to be computed whereas recourse problems involve the expected value of the optimal value function of linear programming problems with random entries.
One way of coping with these difficulties is to employ bounds on probabilities and expected values in algorithms for SLP problems. To make sense these bounds should be more easily computable than the original integrals. We will consider the use of Boole-Bonferroni type bounds for chance constrained problems and the Jensen lower bound along with Edmundson–Madansky type upper bounds for recourse problems.
For chance constrained problems the algorithms are based on general NLP algorithms which are adapted to the special structure. The bounds are used here within a linesearch procedure. Three algorithms will be considered: a reduced gradient method, a cutting plane method and a central cutting plane method. For recourse problems we discuss successive approximation schemes. In these methods lower and upper approximating LP problems are constructed to the SLP problem providing lower and upper bounds for the optimal objective value, respectively. These approximations are refined successively resulting in a convergent successive approximation algorithm.
In the first part of the lecture we summarize the main features of the SLP model types and the above mentioned bounds. The second part is devoted to discussing the usage of the bounds in algorithm design for SLP problems. In the final part of the lecture we present computational results to illustrate the usage of bounds in SLP algorithms. The computer experimentation is based on our model management system SLP–IOR and on our implementation of the algorithms.

Mockus, Jonas
Institute of mathematics and informatics

Bayesian heurisic approach to global and discrete optimization

The traditional numerical analysis considers the optimization algorithms which guarantee some accuracy for all functions to be optimized. This includes the exact algorithms. That is the worst case analysis. Limiting the maximal error requires a computational effort that in many cases increases exponentially with the size of the problem. The alternative is average case analysis where the average error is made as small as possible. The average is taken over a set of functions to be optimized. The average case analysis is called the Bayesian Approach (BA) [2, 7].
There are several ways of applying the BA in optimization. The Direct Bayesian Approach (DBA) is defined by fixing a prior distribution P on a set of functions $f(x)$ and by minimizing the Bayesian risk function $R(x)$ [1, 7]. The risk function describes the average deviation from the global minimum. The distribution P is regarded as a stochastic model of $f(x),\ x \in R^m$ where $f(x)$ might be a deterministic or a stochastic function. In the Gaussian case assuming (see [7]) that the $(n+1)$th observation is the last one

$$R(x) = \frac{1}{\sqrt{2}\pi s_n(x)}$$

$$\int_{-\infty}^{+\infty} \min(c_n, z) e^{-\frac{1}{2}(\frac{y-m_n(x)}{s_n(x)})^2} dz, \tag{1}$$

Here $c_n = \min_i z_i - \epsilon$, $z_i = f(x_i)$, $m_n(x)$ is the conditional expectation given the values of z_i, $i = 1, ..., n$, $d_n(x)$ is the conditional variance, and $\epsilon > 0$ is a correction parameter.
The objective of DBA (used mainly in continuous cases) is to provide as small average error as possible while keeping the convergence conditions.
The Bayesian Heuristic Approach (BHA) means fixing a prior distribution P on a set of functions $f_K(x)$ defining the best values obtained using K times some heuristic $h(x)$ to optimize a function $v(y)$ of variables $y \in R^n$ [8]. As usual the components of y are discrete variables. The heuristic $h(x)$ defines an expert opinion about the decision priorities. It is assumed that the heuristics or their "mixture" depend on some continuous parameters $x \in R^m$, where $m < n$.
The main objective of BHA is improving any given heuristic by defining the best parameters and/or the best "mixtures" of different heuristics. The heuristic decision rules mixed and adapted by BHA often outperform (in terms of speed) even the best individual heuristics as judged by the considered examples. In addition BHA provides almost sure convergence. However, the final results of BHA depend on the quality of the specific heuristics including the expert knowledge. That means the BHA should be regarded as a tool for enhancing the heuristics but not for replacing them.
Many well known optimization algorithms such as Genetic Algorithms (GA) [4], GRASP [6], and Tabu Search (TS) [3], may be regarded as generalized heuristics that can be improved using BHA. There are a number of heuristics tailored to fit specific problems. For example, the Gupta heuristic was the best one while applying BHA to the flow-shop problem [8].

Literatur

[1] DeGroot, M.: *Optimal Statistical Decisions*, McGraw-Hill, New York, 1970.

[2] Diaconis, P.: 'Bayesian Numerical Analysis', *Statistical Decision Theory and Related Topics.* Springer Verlag, 1988, pp. 163–175.

[3] Glover, F.: 'Tabu search: improved solution alternatives', *Mathematical Programming. State of the Art 1994.* University of Michigan, 1994, pp. 64–92.

[4] Goldberg, D. E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989.

[5] Kushner, H.J.: 'A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise', *J. of Basic Engineering* **86** (1964), 97–100.

[6] Mavridou, T., Pardalos, P.M., Pitsoulis, L.S., and Resende, M.G.C.: A GRASP for the biquadratic assignment problem, Tech. rep., AT&T Bell Laboratories, Murray Hill, NJ, 1995.

[7] Mockus, J.: *Bayesian approach to global optimization*, Kluwer Academic Publishers, Dordrecht-London-Boston, 1989.

[8] Mockus, J., Eddy, W., Mockus, A., Mockus, L., and Reklaitis, G.: *Bayesian Heuristic Approach to Discrete and Global Optimization*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

Nollau , Volker (mit D. Hudak)
TU Dresden, Institut für Mathematische Stochastik

Ein modifizierter Gaua-Seidel-Algorithmus mit Suboptimalitätstest für stetig diskontierte semi-Markovsche Entscheidungsprobleme

Es wird gezeigt, daa ein modifizierter Gaua-Seidel-Algorithmus mit Suboptimalitätstest zur näherungsweisen Lösung eines stetig diskontierten semi-Markovschen Entscheidungsproblemes genutzt werden kann. Darüber hinaus wird bewiesen, daa der zustands- und aktionsabhängige Überrelaxationsfaktor dieses Algorithmus gegenüber dem in [1] eingeführten Faktor verbessert werden kann. Ein Beispiel illustriert diese Methode. Falls die Verweilzeiten konstant sind, und der Überrelaxationsfaktor unabhängig vom Zustand und gewählter Aktion ist, erhält man die Resultate von Reetz [2] für diskret diskontierte markovsche Entscheidungsprobleme.

[1] Nollau, V.: Solution of a Discounted Semi-Markovian Decision Problem by Successive Overrelaxation, Optimization (1997), 39, 85 - 97
[2] Reetz, D.: A Decision Exclusion Algorithm for a Class of Markovian Decision Processes, Zeitschrift f. Oper. Res., (1976), 20, 125 - 131

Norkin, Vladimir
Glushkov Institute of Cybernetics, Kiev, Ukraine

Monte Carlo Optimization and Nonstationary Laws of Large Numbers

The estimation of the minimum value F^* of the integral

$$F(x) = \int f(x,\theta)\mathbf{P}(x,d\theta), \quad x \in X \subset R^n,$$

where function $f(x,\theta)$ and measure $\mathbf{P}(x,d\theta)$ are not known explicitly, can be viewed as a significant generalization of the standard Monte Carlo estimation problem. A stochastic approximation procedure produces estimates x^k, $k = 1, 2, \ldots$, of optimal solution x^* by using samples θ^k from distribution $\mathbf{P}(x^k, \cdot)$. An important question is whether

$$F^k = k^{-1} \sum_{s=1}^{k} f(x^s, \theta^s) \longrightarrow F^*, \quad k \longrightarrow \infty, \tag{1}$$

with probability 1, i.e. whether F^* can be estimated from available path dependent observations $\zeta^s = f(x^s, \theta^s)$, $s = 1, 2, \ldots$. The assertion (1) can be derived from the following nonstationary law of large numbers.

Let $\mathcal{F}_k$ be a flow of nondecreasing σ-algebras, random variables ζ^k are $\mathcal{F}_k$-measurable, $k = 1, 2, \ldots$. Denote $z^k = \mathbf{E}\{\zeta^k | \mathcal{F}_{k-1}\}$ conditional mean value of ζ^k with respect to $\mathcal{F}_{k-1}$. Suppose that $\mathbf{E}||\zeta^k - z^k||^{1+\epsilon} < \infty$, $0 < \epsilon \leq 1$. Then with probability 1 $\lim_k \frac{1}{k} \sum_{s=1}^{k} (\zeta^s - z^s) = 0$.

The proof of this and similar results rests on the following stochastic variant of the second Lyapunov's method.

Let $v_k \geq 0, \sigma_k \geq 0, \gamma_k, w_k, k \geq 1$, be random variables. Suppose that each of the following conditions is fulfilled with probability 1:

$$v_{k+1} \leq v_k - \sigma_k w_k + \gamma_k, \quad k \geq 1;$$

$$\lim_k \sigma_k = 0, \quad \sum_{k=1}^{\infty} \sigma_k = +\infty; \quad v_1 + \sum_{k=1}^{\infty} \gamma_k < +\infty.$$

$$\text{If } \liminf_s v_{k_s} > 0, \text{ then } \liminf_s w_{k_s} > 0;$$

$$\text{If } \limsup_s v_{k_s} < \infty, \text{ then } \limsup_s |w_{k_s}| < \infty.$$

Then $\lim_k v_k = 0$ with probability 1.
We discuss various applications of nonstationary laws of large numbers to adaptive Monte Carlo method, Monte Carlo optimization problem, stochastic branch and bound procedure, minimization of risk functions.

Pervukhina , Elena
Sevastopol State Technical University, Sevastopol, Ukraine, Department of Technical Mechanics and Machines

On using differentiation to stochastic approximation algorithm designing

Optimization algorithms differ in their speeds of estimate locating, computational costs, computer-memory requirements and robustness. Real-time algorithms are not always identical to a priori information about noises and characteristic field of dynamic system parameters. That is why they not always lead to well-founded and reliable results in practice.
In the present investigation the concept of Tsypkin (1984) informative approach to the problem of optimal identification algorithm designing is extended. This approach is emphasized in regard to the other approaches, since it is motivated by special choice of gian matrix that provides the more rapid convergence of stochastic approximation algorithm.
The dynamic system under consideration is supposed to be represented by linear difference equation. The problem is to define a parameter vector which minimizes functional of mean losses

$$J(c) = E\{F[e(i), c]\},$$

where E is the mathematical expectation symbol, c is the vector of parameters end $e(i)$ is the vector of identification errors.
To solve the problem stochastic approximation algorithm containing current information from observations is used.
The identification accuracy and convergence of the algorithm are characterized by error covariance matrix:

$$V = E\{(c(n) - c^*)(c(n) - c^*)\}.$$

The stochastic approximation algorithm is optimum as to convergence velocity for gane matrix $B = B^*$ when corresponding error covariance matrix has minimum, that is $V(B) > V(B^*)$ for any $B > 0$. Tsypkin (1984) has determinated matrix B^* by variation method. But as we think this way is bulky and not enough strict.
In the present investigation the strict mathematical substantiation using matrix differential calculus from Magnus (1988) is made. Properties of corresponding algorithm are considered.
Principal results of the present investigation are in full agreement with Tsypkin (1984).

References:
Magnus & Neudecer. Matrix Differential Calculus with Applications in Statistics and Econometrics. Wiley, New York, 1988.
Tsypkin. Foundations of Informational Identification Theory. Nauka, Moscow, 1984.

Qu, Shihong; Marti , Kurt
UniBw München, Fak. LRT

Adaptive Stochastische Bahnplanung in Echtzeit

Das Bahnplanungsproblem für Roboter kann als Variationsproblem dargestellt werden. In der Praxis ist der Modellparameter p des Roboters nicht fest vorgegeben, sondern p muß als Zufallsvektor mit einer bestimmten Verteilung betrachtet werden.
Um eine gegenüber Parameterschwankungen robuste Lösung zu bestimmen, verwenden wir die Methoden der Stochastischen Optimierung und betrachten dazu ein geeignetes Ersatzproblem.
Da man weitere Informationen über den Parameter p sehr oft durch on-line Schätzung bzw. Beobachtung bekommt, ergibt sich das Problem der adaptiven stochastischen Bahnplanung.
Das Adaptive Stochastische Bahnplanungsproblem kann man mittels Finite-Element-Methoden und Parameteroptimierungsverfahren lösen. Wegen des hohen Zeitaufwandes ist diese Vorgehensweise aber nicht in Echtzeit einsetzbar. Durch Approximation mittels neuronaler Netze ergeben sich aber echtzeitfähige adaptive stochastische Bahnplanungsmethoden.
Im Vortrag wird das Verfahren genauer beschrieben und einige numerische Ergebnisse werden präsentiert.

Rackwitz , Ruediger **Hauptvortrag**
Technische Universitaet Muenchen, Institut für Tragwerksbau

Neuere Entwicklungen in der Theorie der Tragwerkszuverlässigkeit

Über den Stand der Berechnung der Tragwerkszuverlässigkeit bei zeitinvarianten und zeitvarianten Problemen wird kurz berichtet. Noch offene Fragen werden diskutiert. Im Hauptteil des Vortrags wird die umgekehrte Fragestellung, wie zu einer gegebenen Zuverlässigkeit optimale Konstruktionsparameter bestimmt werden können, bzw. wie für gegebene Restriktionen die Zuverlässigkeit optimiert werden kann, diskutiert. Hierzu liegen im wesentlichen drei Ansätze vor: ein zweistufiger Algorithmus, die semi-infinite Programmierung und ein einstufiger Algorithmus, der auf der Zuverlässigkeitsmethode erster Ordnung (FORM) aufbaut. Die verschiedenen Ansätze werden im Hinblick auf Zuverlässigkeit und Effizienz miteinander verglichen. Dabei stellt sich heraus, daß der zweistufige Algorithmus, für den noch kein Konvergenzbeweis vorliegt, in der Tat besonderer problemabhängiger Eingriffe bedarf um zu konvergieren. Außerdem ist diese Methode sehr rechenintensiv. Robuster und mit um etwa eine Größenordnung geringerem Aufwand kann man Probleme mit den beiden anderen Methoden lösen. Nur die dritte Methode erlaubt auch eine Optimierung der Zuverlässigkeit. Insgesamt erweist sich, daß der Aufwand bei Optimierung unter Zuverlässigkeitsrestriktionen um etwa eine Größenordnung höher ist als die einfache Zuverlässigkeitsanalyse. Dabei kommt es entscheidend darauf an, ob die Zielfunktion und/oder die Restriktionen die erwarteten Versagenskosten enthalten oder nicht.

Schuëller, G.I. ; Székely, G.S.
Institute of Engineering Mechanics, Leopold-Franzens University, Austria

Reliability Based Optimization with COSSAN

In recent years, the fast development of computational facilities (hardware) and the development of respective methods and procedures allowed the implementation of programs which enable to take into account uncertainties in loading and structural properties for the analysis of structures. When describing these uncertainties in a probabilistic manner, it is possible to quantify the reliability of structures. But often the problems stated in engineering practice are not only to estimate the reliability, but also to build structures at a minimum cost with the constraint of a target reliability. As an example of a program that can be used for reliability based optimization, COSSAN^{TM} (*Co*mputational *S*tochastic *S*tructural *An*alysis) is presented in this workshop. COSSAN^{TM} is toolbox for stochastic structural analysis, programmed in a modular way where all modules can communicate with each other easily. Some of the main module groups in COSSAN^{TM} are:

- Modules for random variable handling and statistics
- Finite Element modules (linear and non-linear)
- Linear Algebra Modules
- Optimization Modules
- Graphical Handling Modules
- etc.

The presentation gives a brief overview on these capabilities of COSSAN^{TM}. The actual performance of the code concentrates on an example from the USER'S MANUAL: "Reliability-Based Frame Design". In this example all the modules mentioned explicitly in the above list are combined to solve a reliability based optimization problem. The problem to solve is to minimize the structural weight of a frame subject to the constraint of a target reliability.

Vietor , Thomas
Ford-Werke AG, Köln

Reliability Oriented Design with Advanced Materials

Design as a separate part of science has gotten bigger importance in the last years. On the one side computer aided design (CAD) is one of the main responsible contributors to this success. On the other side the coupling between CAD and computer based calculation methods (CAE) is responsible. Last but not least the transfer to manufacturing (CIM). Only with increased computers the chain CAD/CAE/CIM can be realized for practical real world problems. This specifically not only for simple single components but increasingly for higher complex products. With increased complexity also the number of demands from different areas increased. There are mainly concurring objectives from different areas like Durability, Safety, NVH (Noise, Vibration and Harshness) and Reliability. For these so called attributes the calculation of objectives and constraints with the help of complicated structural analysis methods is necessary. With increasing demands to new products the necessity for the optimal use of materials will gain importance. For this partially new materials will be developed (fibre reinforced materials, ceramics, rubber materials). Partially the advanced use of old known materials (cast iron, aluminium materials) is possible if new technologies can improve their potential. In the presentation the expression advanced materials will be defined and some representatives from this group will be described. For the evaluation of optimal results the introduction of the specific material laws in the design process as early as possible is necessary. For given reliability constraints deterministic models are often unsufficient. A substitution by stochastic models can be necessary. With the help of examples it will be shown how to introduce different groups of materials early to the design process. One major item is the use of stochastic material properties. With the help of modern CAD/CAE methods development tools can be derived for the practical use. One big area for application is the automotive industry. A number of stochastic quantities have big importance. With increasing cost pressure the necessity of the use of stochastic procedures in the design is mandatory.

Vogel, Silvia
Technische Universität Ilmenau

Stochastische Approximation von Vektoroptimierungsproblemen - mit einer Anwendung auf das Markowitz-Modell der Portfolio-Optimierung

Häufig ist ein Entscheidungsträger mit Optimierungsproblemen konfrontiert, die unbekannte Größen enthalten. In der Regel wird er dann diese Größen schätzen und die so erhaltenen Ersatzprobleme lösen. Es erhebt sich also die Frage nach Eigenschaften der Ausgangsprobleme und der Schätzungen, die sichern, daß die Lösungen der Ersatzprobleme die Lösung des Originalproblems in geeigneter Weise approximieren.
Der Vortrag will diese Problematik für Vektoroptimierungsprobleme betrachten. der Lösungsmengen und von Mengen ϵ-optimaler Punkte diskutiert. Weiterhin sollen stochastische Versionen der 'Ordnungs-Halbstetigkeit' der Effizienzmengen untersucht werden. Zur Illustration der Ergebnisse dient das Markowitz-Modell der Portfolio-Optimierung, wobei auch Shortfall-Restriktionen einbezogen werden.