Asociación Argentina



de Mecánica Computacional

Mecánica Computacional Vol XXIX, págs. 683-697 (artículo completo) Eduardo Dvorkin, Marcela Goldschmit, Mario Storti (Eds.) Buenos Aires, Argentina, 15-18 Noviembre 2010

DESENVOLVIMENTO DE ALGORITMOS MATEMÁTICOS APLICADOS A CONFIABILIDADE ESTRUTURAL

Solange R. dos Santos^{a,b} e Luiz C. Matioli^{b,c}

^aFaculdade Estadual de Ciências e Letras de Campo Mourão, Av. Comendador Norberto Marcondes 733, Cep 87302-060, Paraná, Brasil, solaregina@gmail.com.br, http://www.fecilcam.br

^bPrograma de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Centro Politécnico - Jardim das Américas, Caixa Postal 19081, Brasil, cesec@cesec.ufpr.br, http://www.cesec.ufpr.br

^cUniversidade Federal do Paraná, Setor de Ciências Exatas, Departamento de Matemática, Centro Politécnico - Jardim das Américas, Caixa Postal 19081, Brasil, matioli@ufpr.br, http://www.ufpr.br/~matioli

Palavras Chave: Confiabilidade Estrutural, FORM, Programação Matemática, Método de Duas Fases.

Resumo. Na engenharia diversos estudos têm sido desenvolvidos com o objetivo de analisar a capacidade de uma estrutura suportar a solicitação imposta durante a sua vida útil. O cálculo da probabilidade de falha de uma estrutura envolve um grande número de variáveis aleatórias e possui um grande custo computacional. Neste sentido, nas últimas quatro décadas a utilização de metodologias probabilísticas para avaliação da confiabilidade estrutural têm tido destaque. É essencial para a obtenção da probabilidade de falha de uma estrutura com base nos métodos analíticos de primeira e segunda ordem, FORM (First Order Reliability Method) e SORM (Second Order Reliability Method) a definição de uma função de estado limite, podendo ser modelada por uma função linear ou não-linear, e a determinação do ponto de projeto, que é obtido por meio da resolução de um problema de otimização, cujo objetivo é determinar o ponto de mínima distância da origem até a superfície de falha. Vários algoritmos foram desenvolvidos para a determinação do ponto de projeto, sendo o mais conhecido na literatura o algoritmo HLRF. Diversas melhorias têm sido feitas, no decorrer dos anos, pois na sua forma original o HLFR mostra-se instável podendo não convergir. Sob o enfoque da Otimização, o problema de convergência do algoritmo HLRF é um importante campo para pesquisas. Neste contexto, esta pesquisa propõe um algoritmo que consiste em duas fases: viabilidade e otimalidade. Na fase de viabilidade o objetivo é obter um ponto que reduz uma medida de inviabilidade e na fase de otimalidade procura-se reduzir o valor da função objetivo em relação ao ponto obtido na fase de viabilidade. Nesse algoritmo as fases são mais independentes em relação a outros e, além disso, como o algoritmo baseia-se na abordagem da função penalidade, não é necessária a manutenção da viabilidade a cada iteração. Estudos têm mostrado que a taxa de convergência do método proposto é da mesma ordem que o método do Gradiente Projetado. O desempenho do método será comparado com os resultados de funções de referência na literatura especializada, mostrando assim a aplicabilidade de tal pesquisa.

1 INTRODUÇÃO

Um dos principais objetivos de um projeto estrutural é a garantia de funcionamento do sistema dentro das especificações estabelecidas, com o máximo nível de segurança, levando em conta as restrições econômicas do projeto. Contudo, o cumprimento de tal propósito, principalmente em estruturas de grande porte, não é uma tarefa simples, uma vez que, durante toda a vida útil os sistemas estruturais estão sujeitos a riscos do não atendimento da finalidade para qual foi projetada, levando assim a ocorrência de falhas. Os riscos a qual uma estrutura está sujeita são denominados de probabilidade de falha e são avaliados pela aplicação dos métodos de análise de confiabilidade estrutural. Assim, o objetivo da análise da confiabilidade estrutural é garantir que a resistência de uma estrutura seja maior que a solicitação imposta por toda a sua vida útil. Na metodologia da Confiabilidade Estrutural as grandezas físicas presentes em um projeto, denominadas de variáveis básicas ou variáveis de projeto, são consideradas variáveis aleatórias que podem ser tratadas por meio de um vetor aleatório,

$$X = (X_1, X_2, \dots, X_n) \tag{1}$$

e a probabilidade de falha de uma estrutura é obtida a partir da avaliação das incertezas inerentes as variáveis de projeto por meio das distribuições de probabilidade dessas variáveis aleatórias. Primeiramente, para a avaliação da probabilidade de falha é necessário estabelecer relações funcionais entre as variáveis básicas do sistema estrutural sob consideração. Matematicamente, essa relação ou função de desempenho pode ser descrita como:

$$g(X) = g(X_1, X_2, \dots, X_n).$$
 (2)

A superfície de falha, denominada também de superfície estado limite ou apenas estado limite, é então definida como g(X) = 0, caracterizada como o limite entre as regiões de segurança e de falha.

Na Figura 1 é representada a função estado limite, g(X) = 0, para o caso bidimensional, ou seja, $X = (X_1, X_2)$, a região de segurança, g(X) > 0, e a região de falha ou insegurança, g(X) < 0.



Figura 1: Superfície de falha

Neste sentido, o interesse principal na análise da segurança de uma estrutura está na possibilidade da função estado limite assumir valores pertencentes ao domínio de falha, g(X) < 0. Sendo assim, a probabilidade de falha é dada pela integral:

$$P_{f} = \int \dots \int_{g(X) < 0} f_{X}(x_{1}, x_{2}, \dots, x_{n}) dx_{1} dx_{2} \dots dx_{n}$$
(3)

em que $f_X(x_1, x_2, ..., x_n)$ é a função densidade de probabilidade (f.d.p) conjunta das variáveis aleatórias básicas e a integração é realizada sobre o domínio de falha, g(X) < 0.

O cálculo da probabilidade de falha por meio da equação (3) é considerado complexo, uma vez que envolve integração múltipla e, além disso, raramente se tem conhecimento da f.d.p conjunta das variáveis aleatórias, sendo esta praticamente impossível de se obter.

Não há metodologias que forneçam solução deste problema, sendo assim, soluções aproximadas têm sido obtidas pela aplicação de métodos de simulação e métodos analíticos que aproximam a função estado limite por meio de superfícies lineares (FORM - First-order reliability methods) ou quadráticas (SORM - Second-order reliability methods), no ponto sob a superfície de falha, com maior densidade de probabilidade, denominado de ponto de projeto.

No contexto da Análise de confiabilidade estrutural a obtenção deste ponto constitui um problema de otimização restrito. Neste sentido, o objetivo deste artigo é apresentar uma metodologia diferenciada para a obtenção da probabilidade de falha. Tal metodologia é baseada no Método de Duas Fases proposto inicialmente por Luenberger (1974) que possui a vantagem de ser um método globalmente convergente e possuir uma taxa de convergência assintótica idêntica ao método do Gradiente Projetado.

A Seção 2 desse artigo apresenta alguns conceitos de confiabilidade estrutural, apresentando o desenvolvimento histórico do FORM e do índice de confiabilidade HLRF. A Seção 3 é destinada a apresentação do Método de Duas Fases, mostrando sua relação entre os Métodos do Gradiente Projetado e Função de Penalidade e, além disso, os procedimentos computacionais para a obtenção do índice de confiabilidade estrutural com base na metodologia proposta. A comparação dos resultados com base em experimentos numéricos é discutida na Seção 4. A Seção 5 é dedicada às conclusões do artigo.

2 PRINCIPAIS CONCEITOS DOS MÉTODOS DE PRIMEIRA ORDEM

Os métodos de primeira ordem utilizam informações do primeiro e segundo momento das variáveis aleatórias. Segundo Haldar e Mahadevan (2000) o desenvolvimento histórico dos métodos de primeira ordem, inicialmente foi com base na elaboração dos métodos: First-order second-moment (FOSM) também conhecido como Mean value first-order second-moment (MV-FOSM) e, em seguida, Advanced first-order second-moment (AFOSM).

2.1 Método de primeira ordem e segundo momento

Considere inicialmente a formulação proposta por Cornell (1969) na qual considera apenas o caso de duas variáveis aleatórias básicas $R \in S$, estatisticamente independentes e normalmente distribuídas, em que R representa a resistência e S à solicitação imposta a estrutura. Neste caso, a função estado limite pode ser modelada como,

$$Z = R - S. \tag{4}$$

Sendo assim, Z também é uma variável aleatória normalmente distribuída, em que,

$$Z \sim N\left(\mu_R - \mu_S, \sqrt{\sigma_R^2 + \sigma_S^2}\right).$$
(5)

A falha ocorre quando R < S, isto é Z < 0. Deste modo,

$$P\left(Z<0\right) = P\left(\frac{Z-\mu_Z}{\sigma_Z} < \frac{0-\mu_Z}{\sigma_Z}\right) = P\left(z<-\frac{\mu_Z}{\sigma_Z}\right) = P\left(z<-\frac{\mu_R-\mu_S}{\sqrt{\sigma_R^2+\sigma_S^2}}\right).$$
 (6)

Seja $z \sim N(0, 1)$ e o índice de confiabilidade β dado por:

$$\beta = \frac{\mu_Z}{\sigma_Z} = \frac{\mu_R - \mu_S}{\sqrt{\sigma_R^2 + \sigma_S^2}}.$$
(7)

Temos que a probabilidade de falha pode ser obtida da relação:

$$P_f = \Phi\left(-\beta\right) = 1 - \Phi\left(\beta\right) \tag{8}$$

em que: $\Phi(\beta) = \int_{-\infty}^{\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$ é a função distribuição normal padrão com média zero e variância unitária.

Considere agora a formulação proposta por Rosenbleuth e Esteva (1972), em que é levado em conta o aspecto físico do projeto, cujas variáveis R e S só podem assumir valores positivos. Nesta formulação R e S podem ser consideradas variáveis aleatórias independentes com distribuição lognormal, ou seja, $R \sim LN(\lambda_R, \xi_R) e S \sim LN(\lambda_S, \xi_S)$, (Haldar e Mahadevan, 2000).

Neste caso a função estado limite pode ser modelada como:

$$Y = \frac{R}{S},$$

$$\ln Y = \ln R - \ln S.$$
(9)

Seja $Z = \ln Y$, então:

$$Z = \ln R - \ln S. \tag{10}$$

Observe que, como R e S são lognormais, por definição $\ln R$ e $\ln S$ são variáveis aleatórias normais. Portanto, Z também é uma variável aleatória normalmente distribuída, em que,

$$Z \sim N\left(\lambda_R - \lambda_S, \sqrt{\xi_R^2 + \xi_S^2}\right).$$
(11)

Deste modo, a probabilidade de falha pode ser definida como:

$$P_f = \Phi\left(-\frac{\lambda_R - \lambda_S}{\sqrt{\xi_R^2 + \xi_S^2}}\right) = 1 - \Phi\left(\frac{\lambda_R - \lambda_S}{\sqrt{\xi_R^2 + \xi_S^2}}\right).$$
(12)

Estas formulações que constituem o método FOSM ou MVFOSM, podem ser generalizadas para problemas com mais de duas variáveis aleatórias, representadas pelo vetor X. Assim, a função estado limite pode ser escrita como:

$$Z = g(X) = g(X_1, X_2, \dots, X_n).$$
(13)

Neste caso, o método baseia-se na aproximação de primeira ordem em série de Taylor da função estado limite linearizada nos valores médios das variáveis aleatórias, de modo que:

$$Z \approx g(\mu_X) + \sum_{i=1}^{n} \frac{\partial g}{\partial x_i} (x_i - \mu_{X_i}).$$
(14)

A aproximação para μ_Z e σ_Z é obtida por meio do cálculo da esperança e da raiz quadrada da variância, respectivamente, de Z por meio da expressão (14), obtendo:

$$\mu_Z \approx g\left(\mu_{X_1}, \mu_{X_2}, \dots, \mu_{X_n}\right) = g\left(\mu_X\right),\tag{15}$$

$$\sigma_Z^2 \approx \sum_{i=1}^n \left(\frac{\partial g}{\partial x_i}\right)^2 V(x_i). \tag{16}$$

Portanto, o índice de confiabilidade é:

$$\beta \approx \frac{\mu_Z}{\sigma_Z}.$$
(17)

No entanto, segundo (Haldar e Mahadevan, 2000) a abordagem do método FOSM apresenta algumas deficiências, ou seja, a linearização da função estado limite nos valores médios das variáveis aleatórias acarreta significativos erros no cálculo da probabilidade de falha quando a função é altamente não linear ou tem grande coeficiente de variação, além disso, o índice de confiabilidade não é constante para diferentes formulações da função estado limite sob as mesmas condições de falha e, sua aplicação é limitada a problemas cujas variáveis aleatórias envolvidas são estatisticamente independentes e normalmente distribuídas.

2.2 Método avançado de primeira ordem e segundo momento

Em 1974, Hasofer e Lind propuseram uma nova abordagem aplicada para variáveis aleatórias normais, denominada de método avançado de primeira ordem e segundo momento (AFOSM), que contorna os erros no cálculo da probabilidade de falha acarretados pela linearização da função estado limite nos valores médios das variáveis aleatórias. Sendo assim, para compreensão do método AFOSM, considere uma equação estado limite com duas variáveis:

$$Z = g(R, S) = R - S = 0.$$
 (18)

Neste método, o índice de confiabilidade é calculado com base no espaço reduzido das variáveis de projeto, R' e S', variáveis aleatórias com média zero e desvio padrão unitário, definidas como:

$$R' = \frac{R - \mu_R}{\sigma_R},$$

$$S' = \frac{S - \mu_s}{\sigma_s}$$
(19)

em que, μ_R e μ_S são as médias, e σ_R e σ_S os desvios padrões das variáveis aleatórias de projeto. A notação de variáveis reduzidas será utilizada no decorrer do artigo.

Assim, no espaço das variáveis reduzidas a equação estado limite (18) é transformada para o espaço reduzido, sendo reescrita da seguinte forma:

$$Z = g(R', S') = \sigma_R R' - \sigma_S S' + \mu_R - \mu_S = 0.$$
 (20)

Na Figura 2, as variáveis $R \in S$ constituem o sistema de coordenadas original e $R' \in S'$ o sistema de coordenadas reduzidas. A Figura 2(a) apresenta a equação estado limite, R - S = 0, a região de segurança, R > S, a região de falha, R < S, e o ponto de projeto (R^*, S^*) no sistema de coordenadas originais. Já a Figura 2(b) mostra o sitema de coordenadas reduzidas obtido da transformação das variáveis aleatórias do sistema de coordenadas originais, a região

de segurança, Z > 0, a região de falha, Z < 0, o ponto de projeto, (R'^*, S'^*) , e os pontos de interseção da equação (20) com os eixos R' e S', dados por $\left[-\frac{(\mu_R-\mu_S)}{\sigma_R}, 0\right]$ e $\left[0, \frac{(\mu_R-\mu_S)}{\sigma_S}\right]$, respectivamente. O objetivo é mostrar que, quanto mais próximo a equação estado limite estiver da origem, no sistema de coordanadas reduzidas, maior a região de falha e, consequentemente, quanto mais distante menor a região de falha. Sendo assim, a posição da equação estado limite com relação a origem do sistema de coordenadas reduzidas define uma medida de confiabilidade do sistema.

Portanto, o índice de confiabilidade é definido como a mínima distância da superfície de falha à origem do sistema de coordenadas reduzidas e, o ponto sobre a equação estado limite mais próximo da origem é denominado de ponto de projeto, sendo denotado por (R^*, S^*) no sistema de coordenadas originais e por (R'^*, S'^*) no sistema de coordenadas reduzidas. Deste



Figura 2: Função estado limite linear

modo, o índice de confiabilidade β pode ser obtido utilizando a fórmula de distância entre ponto (origem) e reta (estado limite), ou seja:

$$\beta_{HL} = \frac{|\sigma_R.0 - \sigma_S.0 + \mu_R - \mu_S|}{\sqrt{\sigma_R^2 + \sigma_S^2}} = \frac{\mu_R - \mu_S}{\sqrt{\sigma_R^2 + \sigma_S^2}}.$$
(21)

Em problemas envolvendo um número maior de variáveis, o ponto de projeto é denotado por X^* no sistema de coordenadas originais e por X'^* no sistema de coordenadas reduzidas e representa a pior combinação das variáveis estocásticas, Haldar e Mahadevan (2000).

Observa-se que a expressão (21) coincide com a expressão (7) sendo que as mesmas foram determinadas a partir de procedimentos totalmente diferentes. Isso indica que os Métodos FOSM e AFOSM fornecem os mesmos resultados quando a equação estado limite é linear e as variáveis R e S são normalmente distribuídas e estatisticamente independentes. Em situações práticas, as funções estados limite são não-lineares. Em tais casos, o cálculo da menor distância se torna um problema de otimização, denominado daqui em diante por (P1), e dado por:

$$\min D = \sqrt{X'^T X'}$$

s.a $g(X') = 0$ (22)

onde X' representa as coordenadas do ponto de projeto a ser estimado no sistema de coordenadas reduzidas. Segundo Shinozuka (1983), Ang e Tang (1984) e Melchers (2001) aplicando o Método dos Multiplicadores de Lagrange em (22) obtêm-se as componentes do ponto de projeto dadas por:

$$X_i'^* = -\alpha_i^* \beta, i = 1, 2, \dots, n$$
(23)

em que:

$$\alpha_i^* = \frac{\left(\frac{\partial g}{\partial X_i^{**}}\right)}{\sqrt{\sum_{i=1}^n \left(\frac{\partial g}{\partial X_i^{**}}\right)^2}}$$
(24)

$$\beta = -\frac{\sum_{i=1}^{n} X_i^{\prime *} \left(\frac{\partial g}{\partial X_i^{\prime *}}\right)}{\sqrt{\sum_{i=1}^{n} \left(\frac{\partial g}{\partial X_i^{\prime *}}\right)^2}}$$
(25)

de maneira que, α_i^* são os cossenos diretores ao longo do eixo X'_i , β é o índice de confiabilidade e as derivadas da função estado limite são avaliadas nas componentes do ponto de projeto no espaço reduzido. Ainda, de acordo com Ang e Tang (1984), considerando uma aproximação de primeira ordem da função estado limite g(X) expandida em série de Taylor centrada no ponto de projeto, dada por:

$$g(X) \approx \sum_{i=1}^{n} \left(X'_i - X'^*_i \right) \frac{\partial g}{\partial X'^*_i}$$
(26)

e calculando a estimativa da média e da variância da função estado limite:

$$\mu_{g(X)} \approx -\sum_{i=1}^{n} X_i^{\prime*} \frac{\partial g}{\partial X_i^{\prime*}}$$
(27)

$$\sigma_{g(X)}^2 \approx \sum_{i=1}^n \left(\frac{\partial g}{\partial X_i^{\prime*}}\right)^2.$$
(28)

obtemos a seguinte relação:

$$\frac{\mu_{g(X)}}{\sigma_{g(X)}} = -\frac{\sum_{i=1}^{n} X_i^{\prime*} \frac{\partial g}{\partial X_i^{\prime*}}}{\sqrt{\sum_{i=1}^{n} \left(\frac{\partial g}{\partial X_i^{\prime*}}\right)^2}}.$$
(29)

Percebe-se uma equivalência entre as equações (25) e (29). Portanto a equação (29) também pode ser definida como a mínima distância da origem a um plano tangente a superfície de falha no ponto de projeto. Neste contexto, a aproximação de primeira ordem da $\mu_{g(X)}$ e $\sigma_{g(X)}$ deve ser avaliada no ponto sobre a superfície de falha. Como mencionado anteriormente, no método FOSM, estas aproximações foram avaliadas nos valores médios das variáveis, ocorrendo erros significativos quando g(X) é não-linear, pois a razão $\frac{\mu_{g(X)}}{\sigma_{g(X)}}$ avaliada nos valores médios pode não ser a mínima distância da superfície de falha não-linear até a origem do sistema de coordenadas reduzidas, já que os valores médios não estão necessariamente sobre a superfície de falha.

2.3 Método avançado de primeira ordem e segundo momento para variáveis não-normais

Os métodos analíticos de confiabilidade apresentados até o momento são utilizados para calcular a probabilidade de falha quando as variáveis de projeto são estatisticamente independentes e normalmente distribuídas, caso contrário, tais métodos não fornecem a informação correta sobre a probabilidade de falha. No entanto, na maioria dos problemas práticos as variáveis envolvidas são correlacionas e, além disso, não são normalmente distribuídas e com função de probabilidade complexa.

Neste contexto, segundo Haldar e Mahadevan (2000) e Ayyub e Haldar (1984), Rackwitz, Fiessler, Chend e Lind aperfeiçoaram o método AFOSM, incluindo informações sobre as distribuições das variáveis aleatórias no algoritmo e transformando as variáveis aleatórias não-normais em normais equivalentes. Segundo Melchers (2001) nos casos em que as variáveis aleatórias são não normais correlacionadas elas podem ser transformadas em variáveis aleatórias normais padrão não correlacionadas por meio da Transformação de Rosenblatt e de Nataf. Este mesmo autor apresenta as condições para a aplicação de cada transformação.

No contexto do AFOSM, a probabilidade de falha tem sido estimada usando dois tipos de aproximação para a função estado limite no ponto de projeto: aproximação de primeira ordem FORM e aproximação de segunda ordem SORM, (Haldar e Mahadevan, 2000). O FORM pode ser empregado quando a função de estado limite for uma função linear das variáveis normais não-correlacionadas ou quando a função de estado limite não-linear for representada por uma aproximação linear com variáveis normais equivalentes, já o SORM considera a curvatura da função de desempenho através de um parâmetro caracterizado pela segunda derivada desta função em relação às variáveis. Portanto, o método proporciona melhorias ao FORM ao considerar a não linearidade da função de desempenho, (Flores et al., 2006).

Segundo Haldar e Mahadevan (2000); Madsen et al. (2006); Wang e Grandhi (1996); Liu e Der Kiureghian (1991) e Melchers (2001) comparados a outros algoritmos de otimização não-linear disponíveis na literatura o FORM requer uma quantidade mínima de cálculo a cada iteração, e utiliza somente os valores da função e de sua derivada. Possui convergência rápida, no entanto, em algumas situações, por exemplo, quando a função estado limite é altamente não linear o método converge lentamente, ou ainda, pode divergir para longe da solução.

Numerosas pesquisas têm sido desenvolvidas nas últimas décadas com o objetivo de contornar tais problemas, envolvendo por exemplo, Simulação de Monte Carlo (Mori e Ellingwood, 1993), Teoria do Caos (Yang, 2010) e Yang et al. (2006), Superfície de Resposta (Faravelli, 1989), Métodos dos Momentos (Zhao e Ono, 2001), Algoritmo Genético (Zhao e Jiang, 1995) e Métodos baseados em programação matemática (Liu e Der Kiureghian, 1991) e (Santosh et al., 2006).

3 MÉTODO DE DUAS FASES

Nesta seção será apresentado um algoritmo de programação não-linear proposto por Luenberger (1974) utilizado na resolução de problemas de programação não-linear restrito. Este algoritmo baseia-se na abordagem do Método de Função de Penalidade e no Método do Gradiente Projetado. Para entendimento do método, considere o seguinte problema (P2) de programação não linear:

$$\min f(x)$$

s.a $h(x) = 0$ (30)
 $g(x) \le 0$

em que $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^m$ e $g : \mathbb{R}^n \to \mathbb{R}^r$. Aplicando um particionamento induzido de g em r_1 e r_2 restrições ativas e inativas, respectivamente, as restrições de desigualdade tornam-se $g_A(x) = 0$ e $g_I(x) < 0$. Assuma que todas as funções são contínuas e duas vezes diferenciáveis em todo domínio suficientemente grande e aberto. O método de Duas Fases apresentado por Luenberger, combina as características do Método do Gradiente Projetado com o baixo esforço computacional por passo do método da Função de Penalidade e, utiliza apenas os valores das funções, de suas derivadas primeiras e não requer suposição de convexidade. Cada passo do método é um procedimento de duas fases para resolver um problema penalizado, de modo que, a primeira fase consiste essencialmente de um passo elaborado para melhorar a viabilidade, e a segunda fase é similar ao passo do Gradiente Projetado. As duas seções a seguir abordam brevemente os métodos do Gradiente Projetado e de Função de Penalidade aplicado ao problema de otimização restrito.

3.1 Método do gradiente projetado

Este método é a generalização do método de Cauchy, de forma que o gradiente negativo é projetado em direção a fronteira da região e a busca é feita ao longo da curva resultante, gerando uma seqüência de pontos viáveis convergindo para a solução do problema de minimização restrito. Possui a desvantagem de necessitar satisfazer continuamente a viabilidade durante o processo de minimização.

Vale observar que uma vantagem do Método de Duas Fases é de possuir a mesma taxa de convergência assintótica do Método do Gradiente Projetado, além de contornar as dificuldades associadas com a manutenção da viabilidade exigindo menos cálculos a cada iteração.

3.2 Método da função penalidade

Uma abordagem popular para resolver (P2) é substituí-lo por um problema irrestrito alternativo, dado por:

$$\min f(x) + \mu \|h(x)\|^2 + \mu \|g^+(x)\|^2$$
(31)

em que $g^+ = g\left(g_1^+, g_2^+, \dots, g_r^+\right), g_i^+ = \max\left(0, g_i(x)\right)$ e μ é uma constante positiva grande.

De acordo com a condição necessária de primeira ordem, seja x_0 uma solução local para o problema (P2) e assumindo que os m gradientes de h e r_1 gradientes de g_A avaliadas em x_0 são linearmente independentes, então existem vetores denominados de multiplicadores de Lagrange $v_0 \in R^m$, $\omega_0 \in R^r$ e $\omega_0 \ge 0$, tais que:

$$l(x, v_0, \omega_0) = f(x) + v_0^T h(x) + \omega_0^T g(x)$$

$$\omega_{0i}^T g_i(x_0) = 0$$
(32)

e a langragiana é estacionária com relação a $x \text{ em } x_0$, ou seja, $\nabla_x l(x_0, v_0, \omega_0) = 0$.

O gradiente e a hessiana da função lagrangiana são, respectivamente:

$$\nabla_{x}l\left(x,v_{0},\omega_{0}\right) = \nabla f\left(x\right) + v_{0}^{T}\nabla h\left(x\right) + \omega_{0}^{T}\nabla g\left(x\right),$$

$$\nabla_{xx}^{2}l\left(x,v_{0},\omega_{0}\right) = \nabla^{2}f\left(x\right) + v_{0}^{T}\nabla^{2}h\left(x\right) + \omega_{0}^{T}\nabla^{2}g\left(x\right).$$
(33)

Luenberger mostrou, sob condições apropriadas, que quando a constante de penalidade $\mu \rightarrow \infty$, a solução da função lagrangiana do problema (P2) denotada por x_{μ} tende para a solução do problema original x_0 . Além disso, quando $\mu \rightarrow \infty$, obtém-se:

$$2\mu h\left(x_{\mu}\right) \to v_{0} e \, 2\mu g^{+}\left(x_{\mu}\right) \to \omega_{0} \tag{34}$$

que são os multiplicadores de Lagrange de (P_2) . E, com base nestas condições são definidos os multiplicadores de Lagrange penalizados:

$$v_{\mu}(x) = 2\mu h(x),$$

 $\omega_{\mu}(x) = 2\mu g^{+}(x).$
(35)

3.3 Descrição do método de duas fases

O Método das Duas Fases é uma técnica para minimizar a função objetivo do problema penalizado:

$$\min q(x) = f(x) + \mu \|h(x)\|^2 + \mu \|g^+(x)\|^2.$$
(36)

A primeira fase consiste na aplicação do Método de Newton à função q, com o objetivo de tentar mover-se para perto da região viável (correspondendo ao contorno zero, ou seja, h(x) = 0 e $g^+(x) = 0$) obtendo um ponto z_k que melhora a viabilidade do problema. O passo dado nesta fase é análogo a correção do passo exigido no Método do Gradiente Projetado. Na segunda fase do método um passo é dado na direção oposta ao gradiente de q em z_k . Vale observar que uma busca linear usual é necessária para a realização desta fase.

A Figura 3 ilustra a aplicação do método para o caso de uma única restrição de igualdade e M é o subespaço tangente definido por:

$$M(x) = \{ y : \nabla h(x) \cdot y = 0, \nabla g^{+}(x) \cdot y = 0 \}.$$
(37)



Figura 3: Método das Duas Fases (Adaptado de Luenberger)

Sabe-se que para a aplicação do Método de Newton é necessário o conhecimento da hessiana de q(x) com relação a $v \in \omega$. Luenberger mostrou que para grandes valores de μ , somente informação de primeira ordem é necessária para esse passo também, basta definir no ponto x_k a seguinte função:

$$p(u) = q\left(x_k + \nabla c \left(x_k\right)^T u\right)$$
(38)

em que c(x) é um vetor de dimensão (m+r), estabelecido a partir de uma notação simplificada do conjunto das restrições do problema de minimização, dado por:

$$c(x) = \begin{bmatrix} h(x) \\ g^{+}(x) \end{bmatrix},$$
(39)

 $\nabla c(x_k)$ é a matriz jacobiana de c(x) com dimensão $(m+r) \times n$ e u é um vetor de dimensão $(m+r) \times 1$.

Calculando o gradiente e a hessiana de (38) em relação a u, temos:

$$\nabla p(u) = \nabla c(x_k) \nabla q\left(x_k + \nabla c(x_k)^T u\right),$$

$$\nabla^2 p(u) = \nabla c(x_k) \nabla^2 q\left(x_k + \nabla c(x_k)^T u\right) \nabla c(x_k)^T$$
(40)

 $\nabla p\left(u
ight)$ e $\nabla^{2}p\left(u
ight)$ têm dimensões $(m+r) \times 1$ e $(m+r) \times (m+r)$, respectivamente.

A direção de Newton que minimiza a função objetivo q a partir de x_k é dada por:

$$d_{k} = -\left[\nabla^{2}q\left(x_{k}\right)\right]^{-1}\nabla q\left(x_{k}\right).$$

$$(41)$$

Sabemos que $q(x_k) = p(0)$. Deste modo, devemos calcular (40) quando u = 0. Assim,

$$\nabla p\left(0\right) = \nabla c\left(x_k\right) \nabla q\left(x_k\right),\tag{42}$$

$$\nabla^2 p\left(0\right) = \nabla c\left(x_k\right) \nabla^2 q\left(x_k\right) \nabla c\left(x_k\right)^T.$$
(43)

Portanto, a direção de Newton é dada por:

$$d_{k} = -\left[\nabla^{2} p\left(0\right)\right]^{-1} \nabla p\left(0\right).$$
(44)

Neste momento, é necessário calcular o gradiente e a hessiana da função q:

$$\nabla q(x) = \nabla f(x) + 2\mu \nabla h(x)^T \nabla h(x) + 2\mu \nabla g^+(x)^T g^+(x), \qquad (45)$$

$$\nabla^{2}q(x) = \nabla^{2}f(x) + 2\mu h(x)^{T} \nabla^{2}h(x) + 2\mu \nabla h(x)^{T} \nabla h(x) + + 2\mu g^{+}(x)^{T} \nabla^{2}g^{+}(x) + 2\mu \nabla g^{+}(x)^{T} \nabla g^{+}(x).$$
(46)

As dimensões de $\nabla q(x)$, $\nabla h(x) \in \nabla g^+(x)$ são $n \times 1$, $m \times n \in r \times n$, respectivamente. De (35), temos:

$$\nabla^{2}q(x) = \nabla^{2}f(x) + v_{\mu}(x)^{T}\nabla^{2}h(x) + 2\mu\nabla h(x)^{T}\nabla h(x) + +\omega_{\mu}(x)^{T}\nabla^{2}g^{+}(x) + 2\mu\nabla g^{+}(x)^{T}\nabla g^{+}(x).$$
(47)

Como $\nabla_{xx}^{2} l(x, v_{\mu}(x), \omega_{\mu}(x)) = \nabla^{2} f(x) + v_{\mu}^{T}(x) \nabla^{2} h(x) + \omega_{\mu}^{T} \nabla^{2} g^{+}(x)$, então a equação (47) pode ser reescrita da seguinte forma:

$$\nabla^{2}q(x) = \nabla_{xx}^{2}l(x, v_{\mu}(x), \omega_{\mu}(x)) + 2\mu\nabla h(x)^{T}\nabla h(x) + 2\mu\nabla g^{+}(x)^{T}\nabla g^{+}(x).$$
(48)

Podemos notar de (44) que a avaliação exata da direção de Newton requer informação de segunda ordem da função objetivo e das restrições. Para grandes valores de μ , contudo, considerando $\nabla_{xx}^2 l(x, v_\mu(x), \omega_\mu(x)) = 0$ a matriz hessiana $\nabla^2 q(x_k)$ pode ser aproximada da seguinte forma:

$$\nabla^2 q\left(x_k\right) \approx 2\mu \nabla h\left(x\right)^T \nabla h\left(x\right) + 2\mu \nabla g^+\left(x\right)^T \nabla g^+\left(x\right).$$
(49)

Substituindo (49) em (43), obtemos:

$$\nabla^2 p\left(0\right) \approx 2\mu \left[\nabla c\left(x_k\right) \nabla c\left(x_k\right)^T\right]^2.$$
(50)

Assumindo que $\nabla c(x_k)$ tem posto completo e substituindo (42) e (50) em (44), obtemos:

$$d_{k} = -\frac{1}{2\mu} \left[\nabla c \left(x_{k} \right) \nabla c \left(x_{k} \right)^{T} \right]^{-2} \nabla c \left(x_{k} \right) \nabla q \left(x_{k} \right).$$
(51)

3.4 Descrição do algoritmo do método de duas fases

De acordo com Luenberger, a implementação adequada de um passo do algoritmo é dada por:

Dado um ponto inicial x_k :

Passo 1: Calcular $d_k = -\frac{1}{2\mu} \left[\nabla c \left(x_k \right) \nabla c \left(x_k \right)^T \right]^{-2} \nabla c \left(x_k \right) \nabla q \left(x_k \right).$

Passo 2: Obter o valor β_k que minimiza $q\left(x_k + \beta_k \nabla c \left(x_k\right)^T d_k\right)$ na direção d_k obtida no passo 1. Usar $\beta_k = 1$ como ponto de busca inicial.

Passo 3: Encontrar o ponto $z_k = x_k + \beta_k \nabla c (x_k)^T d_k$ que melhora a viabilidade do problema de minimização.

Os passos de 1 a 3 e 4 a 5 constituem as fase 1 e 2, respectivamente, do método de Duas Fases.

Passo 4: Calcular o valor α_k , aplicando uma técnica de busca linear, que minimiza a função $q(x_k - \alpha_k \nabla q(x_k))$. Usar $\alpha_k = 1$ como ponto de busca inicial.

Passo 5: Obter o ponto $x_{k+1} = z_k - \alpha_k \nabla q(z_k)$ que melhora a otimalidade do problema de minimização.

Passo 6: Fazer $x_k = x_{k+1}$ e reiniciar o algoritmo até que um critério de parada adequado seja satisfeito e a solução do problema seja obtida.

Não é necessário encontrar β_k para garantir boas propriedades de convergência global do método, o que torna a implementação desse algoritmo relativamente simples. Já no caso do método do Gradiente Projetado não há garantia de convergência global sem a introdução de métodos de perturbação, e, além disso, é necessário que o retorno ao conjunto viável, etapa que constitui a primeira fase do processo, seja executado dentro de uma tolerância aceitável. Neste sentido, o baixo esforço por passo no Método das Duas Fases e a quase idêntica taxa de convergência com o método do Gradiente Projetado, indicam a vantagem desse método, (Luenberger, 1974).

3.5 Cálculo da probabilidade de falha

Como mencionado na Seção 2.2, para a obtenção da probabilidade de falha de um sistema é necessário a resolução de (P1), que é um caso particular do problema proposto por Luenberger, considerando apenas restrições de igualdade. Vamos agora definir (P3) como sendo o problema de calcular a probabilidade de falha considerando apenas uma função estado limite h(x) = 0, da seguinte forma:

$$\min f(x) = \sqrt{x^T x}$$

s.a $h(x) = 0$ (52)

em que $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ e $h : \mathbb{R}^n \to \mathbb{R}$. Lembrando que, x é um vetor aleatório cujas componentes são as variáveis de projeto no espaço normal padronizado. O método de Duas Fases pode ser aplicado para resolver o seguinte problema:

$$\min q(x) = f(x) + \mu \left[h(x) \right]^2$$
(53)

sendo q(x) a função penalizada para o problema (P3). Nestes termos, o algoritmo pode ser estabelecido da seguinte forma:

Dado um ponto inicial x_k :

Passo 1: Calcular
$$d_k = -\frac{1}{2\mu} \left[\nabla h \left(x_k \right)^T \nabla h \left(x_k \right) \right]^{-2} \nabla h(x_k)^T \nabla q \left(x_k \right)^T$$

Passo 2: Obter o valor β_k que minimiza $q(x_k + \beta_k \nabla h(x_k) d_k)$ na direção d_k obtida no passo 1.

Passo 3: Encontrar o ponto $z_k = x_k + \beta_k \nabla h(x_k) d_k$.

Passo 4: Calcular o valor α_k , aplicando uma técnica de busca linear, que minimiza a função $q(x_k - \alpha_k \nabla q(x_k))$.

Passo 5: Obter o ponto $x_{k+1} = z_k - \alpha_k \nabla q(z_k)$.

Passo 6: Fazer $x_k = x_{k+1}$ e reiniciar o algoritmo até obtenção da solução.

4 EXPERIMENTOS NUMÉRICOS

Com o objetivo de demonstrar a validade e aplicabilidade do método proposto, o algoritmo mencionado anteriormente foi implementado em Matlab 7.8, versão R2009a, em um computador com processador Intel(R) Core(TM)2 Duo CPU T5870 2,00GHz e 3,00GB de memória RAM. Vários testes numéricos foram realizados e alguns resultados são apresentados resumidamente nas tabelas a seguir. A Tabela 1 apresenta dez problemas, as funções estados limite retiradas da literatura especializada e que por sua vez, são utilizadas com o objetivo de analisar a eficácia da metodologia proposta. Já na Tabela 2 é apresentada a descrição das variáveis aleatórias envolvidas no problema, ou seja, a distribuição de probabilidade de cada variável de projeto com seus respectivos parâmetros, média e variância. E, por fim a Tabela 3 apresenta, para os dez problemas testados, uma comparação entre o índice de confiabilidade β obtido por meio da metodologia proposta e os resultados observados na literatura, que estão referenciados na última coluna da referida tabela.

Problema	Função estado limite		
1	$h(x) = x_1 x_2 - 146, 14$		
2	$h(x) = 2 + 0.015 \left(\sum_{i=1}^{9} x_i^2\right) - x_{10}$		
3	$h(x) = 0, 1(x_1 - x_2)^2 - \frac{(x_1 + x_2)}{\sqrt{2}} + 2, 5$		
4	$h(x) = -0, 5(x_1 - x_2)^2 - \frac{(x_1 + x_2)}{\sqrt{2}} + 3$		
5	$h(x) = 2 - x_2 - 0, 1x_1^2 + 0, 06x_1^3$		
6	$h(x) = 2,5 - 0,2357(x_1 - x_2) + 0,0046(x_1 + x_2 - 20)^4$		
7	$h(x) = 3 - x_2 - 256x_1^4$		
8	$h(x) = x_1^3 + x_2^3 - 18$		
9	$h(x) = x_1^3 + x_2^3 - 18$		
10	$h(x) = x_1^3 + x_2^3 - 67,5$		

Tabela 1: Funções estados limite.

Com base nos resultados obtidos para os problemas apresentados na Tabela 1, pode-se notar que o algoritmo apresentou consistência no cálculo da probabilidade de falha quando comparados aos resultados obtidos na literatura, como pode ser visto na Tabela 3. Neste contexto, o emprego do Método de Duas Fases em problemas de confiabilidade estrutural é uma área de pesquisa promissora. Sendo assim, uma possibilidade de aprimoramento de tal metodologia é a inserção de uma busca linear no passo de Newton, que deverá apresentar bons resultados para problemas que envolvam maior número de variáveis aleatórias e também variáveis aleatórias não-normais.

Problema	Descrição das variáveis		
1	$x_2 \sim N(78064, 4; 11709, 7^2)$ e $x_2 \sim N(0, 0104, 4; 0, 00156^2)$		
2	$x_{110} \sim N(0, 1)$		
3, 4, 5 e 7	$x_{1,2} \sim N(0,1)$		
6	$x_{1,2} \sim N(10, 3^2)$		
8	$x_{1,2} \sim N(10, 5^2)$		
9 e 10	$x_1 \sim N(10, 5^2) \text{ e } x_2 \sim N(9, 9; 5^2)$		

Tabela 2: Distribuição de probabilidade das variáveis aleatórias de projeto

Problema	β - Metodologia proposta	β - Literatura
1	5,3653	5,443 (Grooteman, 2008)
2	2,0000	2,103 (Grooteman, 2008)
3	2,5000	2,4810 (Grooteman, 2008)
		2,5000 (Borri e Speranzini, 1997)
4	1,6583	1,6250 (Grooteman, 2008)
		1,6580 (Borri e Speranzini, 1997)
5	1,9998	1,9960 (Grooteman, 2008)
6	2,5000	2,4310 (Grooteman, 2008)
7	3,0000	2,925 (Grooteman, 2008)
8	2,2401	2,24009 (Santosh et al., 2006)
		2,2401 (Yu et al., 1997)
9	2,2260	2,2594 (Santosh et al., 2006)
		2,2260 (Yu et al., 1997)
10	1,9003	1,90026 (Santosh et al., 2006)
		1,9 (Tu e Choi, 1997)

Tabela 3: Comparação dos resultados

5 CONCLUSÃO

Neste artigo, apresentamos uma nova metodologia para o cálculo do índice de confiabilidade estrutural. A metodologia é baseada no método de Duas Fases proposto no artigo de Luenberger (1974), para um problema de programação não linear restrito. A utilização deste método em problemas de confiabilidade é motivada por ser um método de simples implementação, requerer apenas informação de primeira ordem das funções envolvidas e não necessitar da suposição de convexidade. A convergência do método é semelhante à do Gradiente Projetado e tem a vantagem em relação a este por não requerer viabilidade nos pontos gerados pelo algoritmo a cada iteração. O algoritmo é promissor, pois a cada iteração, faz apenas um passo de Newton e um passo de Cauchy, sendo que a direção de Newton (relaxada) é determinada utilizando apenas o gradiente da função, além disso, o algoritmo é muito simples, facilmente implementável e pelos testes realizados até o momento se mostrou eficiente.

REFERÊNCIAS

Ang A. e Tang W.H. *Probability concepts in engineering planning and design*, volume II. John Wiley & Sons, 1984.

Ayyub B. e Haldar A. Practical structural reliability techniques. *Journal of Structural Engineering*, 110:1707–1724, 1984.

- Borri A. e Speranzini E. Structural reliability analysis using a standard deterministic finite element code. *Structural Safety*, 19:361–382, 1997.
- Faravelli L. Response-surface approach for reliability analysis. *Journal of Engineering Mechanics*, 115:2763–2781, 1989.
- Flores J., Viana F., e Rade D. Métodos de análise de confiabilidade de primeira e segunda ordem aplicados a sistemas estruturais. *VII Simpósio de Mecânica Computacional*, 2006.
- Grooteman F. Adaptive radial-based importance sampling method for structural reliability. *Structural Safety*, 30:533–542, 2008.
- Haldar A. e Mahadevan S. *Probability, reliability and statistical methods in engineering design.* John Wiley & Sons, 2000.
- Liu P.L. e Der Kiureghian A. Optimization algorithms for structural reliability. *Structural Safety*, 9:161–177, 1991.
- Luenberger D.G. A combined penalty function and gradient projection method of nonlinear programming. *Journal of Optimization Theory and Applications*, 14:477–495, 1974.
- Madsen H., Krenk S., e Lind N. Methods of structural safety. Dover publications, 2006.
- Melchers R. Structural reliability analysis and prediction. John Wiley & Sons, 2001.
- Mori Y. e Ellingwood B. Time-dependent system reliability analysis by adaptive importance sampling. *Structural Safety*, 12:59–73, 1993.
- Santosh T., Saraf R., Ghosh A., e Kushwaka H. Optimum step length selection rule in modified HL-RF method for structural reliability. *International Journal of Pressure Vessels and Piping*, 83:742–748, 2006.
- Shinozuka M. Basic analysis of structural safety. *Journal of Structural Engineering*, 109:721–740, 1983.
- Tu J. e Choi K. A performance measure approach in reliability based structural optimization. *Technical Report R97-02: The University of Iowa*, 1997.
- Wang L. e Grandhi R. Safety index calculations using intervening variables for structural reliability. *Computers and Structures*, 59:1139–1148, 1996.
- Yang D. Chaos control for numerical instability of first order reliability method. *Commun Nonlinear Sci Numer Simulat*, 15:3131–3141, 2010.
- Yang D., Li C., e Cheng G. Convergence analysis of first order reliability method using chaos theory. *Computers and Structures*, 84:563–571, 2006.
- Yu X. C., K.K., e Chang K. Reliability and durability based design sensitivity analysis and optimization. *Technical Report R97-02: The University of Iowa*, 1997.
- Zhao Y. e Jiang J. A structural reliability method basead on genetic algorithm. *Earthquake Engineering and Engineering Vibration*, 15:44–58, 1995.
- Zhao Y. e Ono T. Moment methods for structural reliabilty. Structural Safety, 23:47-75, 2001.