

Príspevok k určeniu konjugovaného gradientu v sústavách automatického riadenia

PAVEL KNEPPO

Predkladaná práca sa zaoberá príkladom použitia metódy citlivosti pre výpočet zložiek konjugovaného gradientu optimalizačného kritéria. Porovnávajú sa výhody optimalizačného postupu získané „klasickou“ gradientovou metódou a metódou konjugovaného gradientu.

ÚVOD

Veľká výhoda použitia gradientových metód pre optimalizáciu riadených sústav spočíva v tom, že nie je potrebné podrobne poznať matematický model danej sústavy, ale stačí poznať len hodnoty zložiek gradientu v daných bodoch. Z toho plynie veľká časová úspora.

Každá optimalizačná metóda je charakterizovaná potrebnou dobou výpočtu alebo pri iteračnom postupe, počtom iteračných krokov, potrebných na dosiahnutie optimálneho stavu. Všeobecne je tento počet závislý od počiatočných podmienok a od voľby kroku.

Z hľadiska riadenia reálnych procesov sú výhodnejšie metódy, ktoré dovoľujú dosiahnutie optimálnej oblasti z daných počiatočných podmienok v čo možno najkratšom čase (s najmenším počtom krokov).

M. R. Hestenes, E. Stiefel [1], R. Fletcher, C. M. Reeves [2], J. W. Daniel [3], L. S. Lasdon, S. K. Mitter, A. D. Waren [4] a ďalší, popísali metódu konjugovaného gradientu ako metódu vhodnú pre riešenie lineárnych sústav a dokázali, že v niektorých prípadoch vedie k výsledkom rýchlejšie než „klasická“ gradientová metóda.

V našom prípade sa budeme zaoberať porovnaním oboch uvedených metód pri riešení úlohy optimalizácie sústavy druhého rádu.

Predkladaná práca sa zaoberá príkladom použitia metódy citlivosti pre výpočet zložiek konjugovaného gradientu optimalizačného kritéria. Porovnávajú sa výhody optimalizačného postupu získaného „klasickou“ gradientovou metódou a metódou konjugovaného gradientu.

1. VÝPOČET ZLOŽIEK VEKTORU $\text{grad } J$ METÓDOU CITLIVOSTI

Nech je sústava S opísaná diferenciálnou rovnicou v nasledujúcom tvare:

$$(1.1) \quad f(x^{(l)}, x^{(l-1)}, \dots, x^{(k)}, \dots, \dot{x}, x, q_1, \dots, q_n, \dots, q_m, t) = 0,$$

158 kde $x^{(k)}$ – k -tá derivácia vstupného parametra, q_i – i -tý parameter sústavy, t – čas, a ďalej o funkcii f sa predpokladá, že je to funkcia spojitá a diferencovateľná podľa parametrov x a q_i ($i = 1, 2, \dots, n$).

Derivovaním diferenciálnej rovnice (1.1) podľa parametru q_i ($i = 1, 2, \dots, n$), obdržíme systém lineárnych diferenciálnych rovníc v tvare :

$$(1.2) \quad \frac{\partial f}{\partial x^{(l)}} \Phi_i^{(l)}(t) + \frac{\partial f}{\partial x^{(l-1)}} \Phi_i^{(l-1)}(t) + \dots + \frac{\partial f}{\partial x^{(k)}} \Phi_i^{(k)}(t) + \dots + \\ + \frac{\partial f}{\partial \dot{x}} \dot{\Phi}_i(t) + \frac{\partial f}{\partial x} \Phi_i(t) = - \frac{\partial f}{\partial q_i} \\ (i = 1, 2, \dots, n),$$

ktorý nazývame systémom rovníc citlivosti, a funkcie $\Phi_i^{(l)}(t)$, $\Phi_i^{(l-1)}(t)$, ..., $\Phi_i^{(k)}(t)$, ..., $\dot{\Phi}_i(t)$, $\Phi_i(t)$ nazývame funkcie citlivosti alebo vplyvové koeficienty, ktoré vypočítame riešením diferenciálnych rovníc (1.1) a (1.2) [5], [6].

Použitie funkcií citlivosti pre (1.1), kde q_1, q_2, \dots, q_n sú konštantné parametre, spočíva v nájdení takých hodnôt parametrov q_i ($i = 1, 2, \dots, n$), ktoré minimalizujú daný funkcionál

$$(1.3) \quad I(q_1, q_2, \dots, q_n) = \int_0^T \Omega[x^{(1)}(q_1, \dots, q_n), x^{(l-1)}(q_1, \dots, q_n), \dots, x^{(k)}(q_1, \dots, q_n), \dots \\ \dots, \dot{x}(q_1, \dots, q_n), x(q_1, \dots, q_n), t] dt,$$

kde Ω je podintegrálna funkcia daného funkcionálu, T je doba procesu prebiehajúceho v riadenej sústave.

Zložky vektoru grad I môžeme písať v tvare parciálnych derivácií ako

$$(1.4) \quad \frac{\partial I}{\partial q_i} = \int_0^T \left[\frac{\partial \Omega}{\partial x^{(l)}} \Phi_i^{(l)}(t) + \frac{\partial \Omega}{\partial x^{(l-1)}} \Phi_i^{(l-1)}(t) + \dots + \right. \\ \left. + \frac{\partial \Omega}{\partial \dot{x}} \dot{\Phi}_i(t) + \frac{\partial \Omega}{\partial x} \Phi_i(t) \right] dt \\ (i = 1, 2, \dots, n),$$

kde $\Phi_i(t) = \partial x / \partial q_i$ sú známe funkcie citlivosti.

Ak napríklad funkcionál (1.3) je tvaru

$$(1.5) \quad I = \int_0^\infty \Omega(\varepsilon, t) dt,$$

kde podintegrálna funkcia Ω môže byť napr. $\Omega_1 = |\varepsilon|$, $\Omega_2 = \varepsilon^2$, ..., ε je odchylka

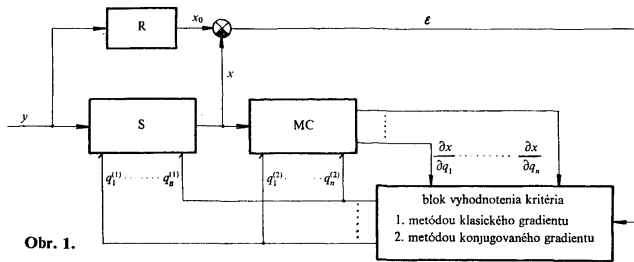
ktorá podľa obr. 1 bude nasledujúceho tvaru

$$(1.6) \quad \varepsilon(t) = x_0(t) - x(t, q_i) \\ (i = 1, 2).$$

Ak teraz derivujeme vzťah (1.6) podľa parametru q_i ($i = 1, 2$), obdržíme, že

$$(1.7) \quad \frac{\partial \varepsilon}{\partial q_i} = - \frac{\partial x}{\partial q_i},$$

kde pravá strana je opäť funkcia citlivosti.



Obr. 1.

Prevedením parciálnej derivácie vzťahu (1.5) podľa parametru q_i ($i = 1, 2$) obdržíme hľadané zložky grad I v nasledujúcom tvare

$$(1.8) \quad \frac{\partial I}{\partial q_i} = \int_0^{\infty} \frac{\partial \Omega}{\partial \varepsilon} \frac{\partial \varepsilon}{\partial q_i} dt \\ (i = 1, 2).$$

Zo vzťahu (1.4) a (1.8) je vidieť, že vypočítať zložky grad I , vedie teda k výpočtu funkcií citlivosti. Vhodná metóda, ktorá rieši túto úlohu na analógovom počítači je metóda bodov citlivosti [5], [6], ktorá je použitá i v predkladanej práci.

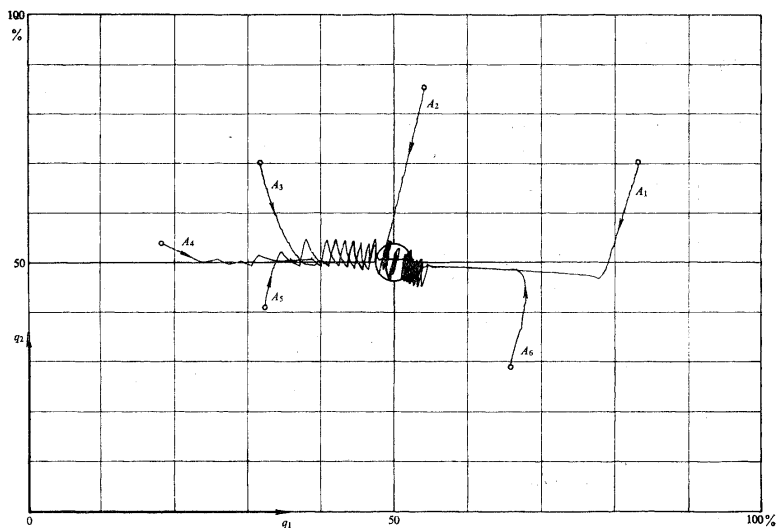
Ako príklad bola vyšetovaná lineárna sústava S s prenosom tvaru

$$(1.9) \quad S(p) = \frac{1}{a_2 p^2 + a_1 p + a_0},$$

pričom integrálne kritérium bolo v tvare integrálu z absolútnej hodnoty odchyľky t.j.

$$(1.10) \quad I = \int_0^T |e| dt, \quad |e| = |x_0 - x|.$$

160 Podľa obr. 1, ktorý predstavuje blokové schéma úlohy značí: S – sústava, R – referenčný model, MC – model citlivosti [5], [6], x – výstupná veličina zo sústavy, x_0 – výstupná veličina z referenčného modelu t.j. žiadaná veličina, y – vstupná veličina v tvare jednotkového skoku.



Obr. 2.

Proces optimalizácie je uskutočňovaný zmenou parametrov q_1, q_2 pomocou modelu citlivosti MC, ktorý má v našom prípade rovnakú štruktúru ako samotná sústava S. Na obr. 2 v rovine parametrov q_1, q_2 sú znázornené trajektórie $A_1 - A_6$ z jednotlivých počiatočných podmienok pri pohybe do optima klasickou metódou gradientu.

Pristúpme ďalej k procesu optimalizácie použitím konjugovaného gradientu.

2. METÓDA KONJUGOVANÉHO GRADIENTU

Ide tu v podstate o určitú modifikáciu, ktorá aby splnila svoj účel, musí viesť k progresívnejšiemu výpočtu než ako tomu je v prípade klasickej metódy gradientu. Ukážme si, že je možné zaviesť takúto úpravu.

Všeobecný algoritmus pre výpočet konjugovaného gradientu v $i + 1$ kroku bude nasledovný [4]

$$(2.1a) \quad \mathbf{p}_0 = -\mathbf{g}_0,$$

$$(2.1b) \quad \mathbf{p}_{i+1} = -\mathbf{g}_{i+1} + \beta_i \mathbf{p}_i,$$

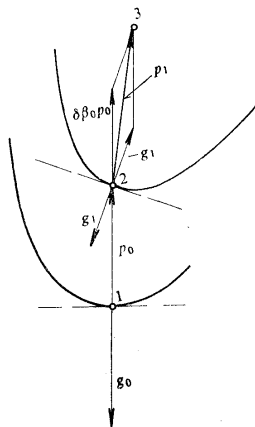
kde \mathbf{q}_0 je počiatočná hodnota parametru, $\mathbf{g}_0 = \mathbf{g}(\mathbf{q})$ pre $\mathbf{q} = \mathbf{q}_0$ a

$$(2.2) \quad \beta_i = \mathbf{g}_{i+1} \mathbf{q}_{i+1} / \mathbf{g}_i \mathbf{g}_i,$$

kde skalárny súčin $\mathbf{g}_i \mathbf{g}_j$ je definovaný ako

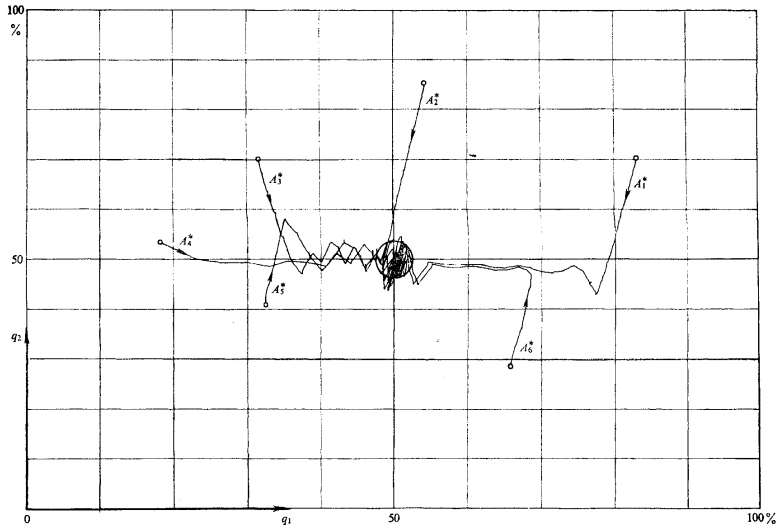
$$(2.3) \quad \mathbf{g}_i \mathbf{g}_j = \sum_{k=1}^{\infty} g_{ik} g_{jk}.$$

Podľa algoritmu (2.1a,b) je vidieť, že nový smer \mathbf{p}_{i+1} nebude len záporne vzatý gradient $-\mathbf{g}_{i+1}$, ale bude doplnený o hodnotu súčinu $\beta_i \mathbf{p}_i$, pričom ako vyplýva zo vzťahu (2.2), β_i je funkciou aj predchádzajúceho kroku. Znamená to teda, že zobrazujúci bod pri prvom kroku pôjde po rovnakej trajektórii pre obidva prípady t.j. je splnený vzťah (2.1a) ale pri druhom kroku sa prejaví určitá „váha“ súčinu $\beta_i \mathbf{p}_i$.

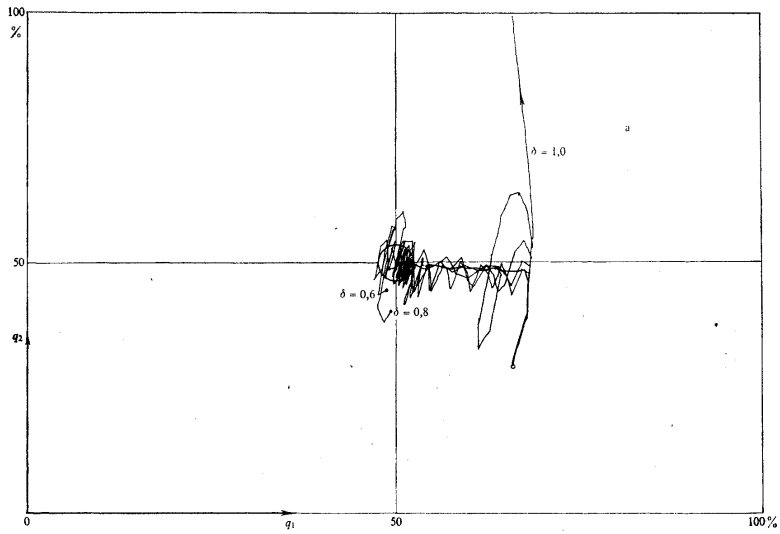


Obr. 3.

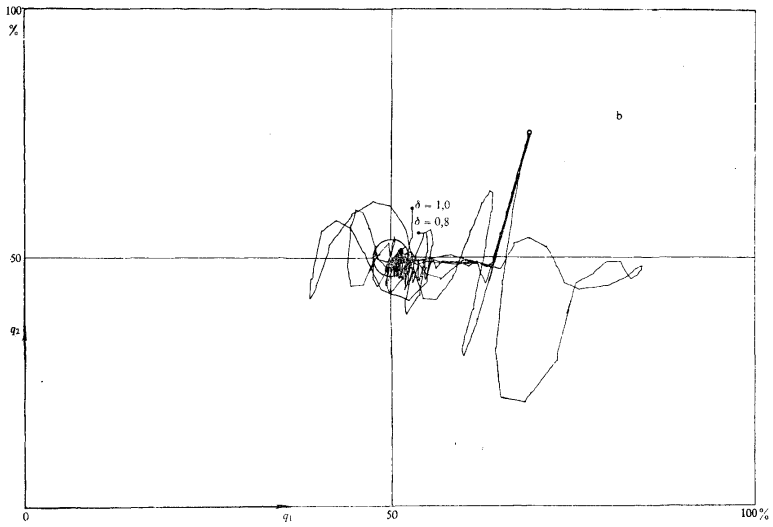
Algoritmus (2.1a,b) vysvetľuje obr. 3. Pohyb do optima v rovine parametrov q_1, q_2 metódou konjugovaného gradientu je znázornený na obr. 4 trajektóriami $A_1^* - A_6^*$. Riešenie je prevádzané na rovnakom príklade (obr. 1) z rovnakých počiatočných podmienok.



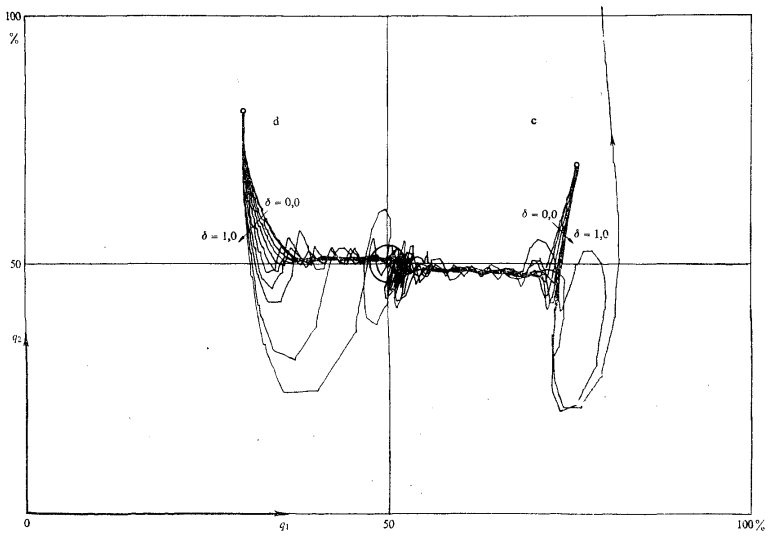
Obr. 4.



Obr. 5.



Obr. 6.

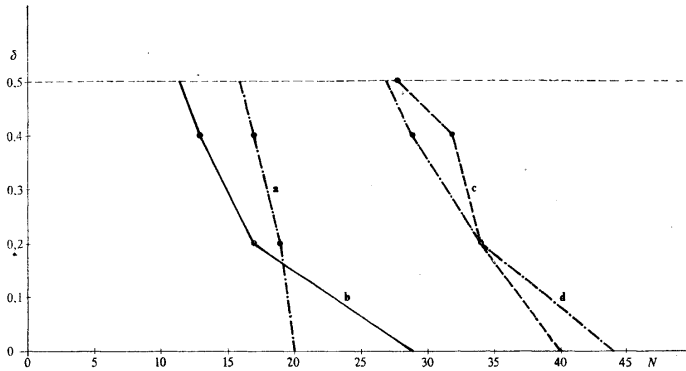


Obr. 7.

L. J. ...

3. POROVNANIE KLASICKEJ A KONJUGOVANEJ METÓDY GRADIENTU

V tabuľke 1 sú zhrnuté výsledky riešenia. V ľavom stĺpci $A_1 - A_6$ označuje príslušné trajektórie podľa obr. 2, t.j. bez konjugovaného gradientu, $A_1^* - A_6^*$ označuje trajektórie s konjugovaným gradientom. Pravý stĺpec číselne vyjadruje o koľko % nastalo zlepšenie $\xi = (N - N^*)/N$, kde N je počet iterácií metódou klasického gra-



Obr. 8.

dientu a N^* je počet iterácií metódou konjugovaného gradientu, k je veľkosť kroku, δ je konštanta. Ako ukázali merania pre náš konkrétny prípad, vzťah (2.1b) bolo potrebné doplniť na tvar

$$(3.1) \quad \mathbf{p}_{i+1} = -\mathbf{g}_{i+1} + \delta\beta_i\mathbf{p}_i.$$

Vo výpočte sme sa obmedzili na ten prípad, aby riešenie bolo ukončené za najmenší počet iteračných krokov a ďalej aby oscilácie okolo optima neprekročili oblasť dovolených oscilácií 5% (na obr. 2, 4 - 7 označené krúžkom). Ako ukazujú merania na obr. 5, 6, 7, kde z náhodne zvolených počiatočných podmienok v rovine parametrov q_1, q_2 bolo realizované riešenie pri rôznych hodnotách δ , vidieť napr. na obr. 5 trajektória „a“, že pri hodnote $\delta = 0,6$ a $\delta = 0,8$ riešenie bolo nevhodné pre veľké oscilácie okolo optima.

Optimálna hodnota, ktorá ešte vyhovuje zadaným podmienkám je $\delta = 0,5$, viz. obr. 8, kde je grafická závislosť počtu iteračných krokov N na hodnote δ .

Ako je vidieť z tabuľky 1, použitím konjugovanej metódy gradientu v porovnaní s klasickou metódou čo do počtu iteračných krokov docielime v niektorých prípadoch viac ako 50% zlepšenie. Priemerná hodnota zlepšenia ξ číni 47,7%.

Tabuľka 1.

Trajektória	Počet iterácií	k	δ	$\xi = (N - N^*)/N [\%]$
A_1 A_1^*	52 22	0,04	0,0 0,4	57,7
A_2 A_2^*	18 8		0,0 0,4	55,6
A_3 A_3^*	22 16		0,0 0,4	27,2
A_4 A_4^*	29 12		0,0 0,4	58,7
A_5 A_5^*	19 10		0,0 0,4	47,4
A_6 A_6^*	23 14		0,0 0,4	39,1

ZÁVER

V citovanej literatúre bola otázka porovnania metódy klasického a konjugovaného gradientu rozpracovaná podrobne a boli jasne ukázané výhody najmä jej časovej úspornosti. Predkladaná práca si kládla za cieľ prispieť praktickým príkladom k tomuto porovnaniu. Na rozdiel od citovaných prác boli k výpočtu použité vlastnosti funkcií citlivosti. Z praktického hľadiska je zlepšenie použitím konjugovanej metódy gradientu významnejšie, než potreba nových obvodoých prvkov v riadiacej sústave.

(Došlo dňa 12. mája 1968.)

LITERATÚRA

- [1] M. R. Hestenes, E. Stiefel: Methods of conjugate gradients for solving linear systems. J. Research NBS 49 (1952), 409.
- [2] R. Fletcher, C. M. Reeves: Function minimization by conjugate gradients. British Computer J. (July 1964), 149—154.
- [3] J. W. Daniel: Convergence of the conjugate gradient method with computationally convenient modifications. Numerische Mathematik 10 (1967), 125—131.
- [4] L. S. Lasdon, S. K. Mitter, A. D. Waren: The conjugate gradient method for optimal control problems. IEEE Trans. on Automatic Control AC-12 (1967), 2, 132—138.

- [5] S. Bingulac, P. Kokotovic: Automatic optimization of linear feedback control systems on an analog computer. Proceedings of the International Assoc. for Analog Computation VII (Jan. 1965,, 1. (Brussels).
- [6] P. Kneppo: Použitie teórie citlivosti v systémoch automatického riadenia. Strojnícky časopis XVIII, (1967), 5.

SUMMARY

A Contribution to the Determination of Conjugated Gradient in Automatic Control Systems

PAVEL KNEPPO

The submitted paper deals with an example of the application of the sensitivity method for the computation of the conjugated gradient components of the optimization criterion. The advantages of the optimizing procedure achieved by the "classic" gradient method and that of the conjugated gradient are compared.

Ing. Pavel Kneppo, Ústav technickej kybernetiky SAV, Dúbravská cesta, Bratislava.