

5.9. Egenvärdesproblem, iterativa lösningsmetoder

Vi skall nu övergå till att studera **egenvärdesproblemet** för en $n \times n$ matris A som kan formuleras på följande sätt: För vilka värden av λ har det homogena ekvationssystemet

$$(A - \lambda I)x = 0$$

en sådan lösning x , att $x_i \neq 0$ åtminstone för ett värde av i ? Enligt teorin för linjära ekvationer är villkoret för en icke-trivial lösning att systemets koefficientdeterminant försvinner, dvs att $\det(A - \lambda I) = 0$.

Härav följer, att egenvärdena λ_i är de n rötterna (*latent* rötterna) till den **karaktäristiska ekvationen** (även kallad *sekulärekvationen*, av historiska skäl)

$$p_A(\lambda) = \det(A - \lambda I) = 0,$$

där $p_A(\lambda)$, det **karaktäristiska polynomet**, är ett n :te gradens polynom i λ . Egenvärdena kan alltså beräknas skilt för sig. Om ett egenvärde λ_i är känt, så är den motsvarande egenvektorn x_i en lösning till ovanstående linjära ekvationssystem.

Om man å andra sidan känner en egenvektor x_i , så kan man omedelbart beräkna det motsvarande egenvärdet ur ekvationen

$$\lambda_i = \frac{x_i^T A x_i}{x_i^T x_i}$$

Uttrycket i högra membrum brukar också kallas **Rayleighs kvot**, då x är en godtycklig vektor. Teoretiskt kan man finna egenvärdena genom att bestämma koefficienterna i det karakteristiska polynomet

$$p_A(\lambda) = (-1)^n \lambda^n + b_1 \lambda^{n-1} + \dots + b_n$$

och sedan lösa $p_A(\lambda) = 0$ algebraiskt.

Även om det utvecklats speciella metoder att göra detta (bl.a. av Danilevsky och Krylov) rekommenderas ett dylikt förfaringssätt dock endast för små matriser med välseparerade egenvärden. Små störningar i koefficienterna b_k kan leda till stora störningar i egenvärdena, även då små störningar i A :s matriselement leder till störningar av samma storleksordning i egenvärdena.

För att lokalisera egenvärdena kan man använda **Gerschgorins cirklar**. Antag, att A är en $n \times n$ matris, och att $C_i, i = 1, 2, \dots, n$ är cirkelskivor i det komplexa planet med medelpunkterna a_{ii} och radierna $R_i = \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|$, dvs definierade genom villkoret $|z - a_{ii}| \leq R_i$. Antag vidare, att $D = \cup_{i=1}^n C_i$.

Då är mängden av alla egenvärden för A en delmängd av D (Gerschgorins sats¹). D är alltså en region i det komplexa planet som innehåller alla egenvärden för den (komplexa) matrisen A . Detta kan visas på följande sätt: Antag, att λ är ett egenvärde för A och x den motsvarande egenvektorn, som är normerad så att $\max_i |x_i| = 1$. Vi skriver egenvärdesekvationen $\lambda x = Ax$ i formen

$$(\lambda - a_{ii})x_i = \sum_{\substack{k=1 \\ k \neq i}}^n a_{ik}x_k, \quad i = 1, 2, \dots, n$$

Om $|x_s| = 1$ är den största komponenten av x , så gäller alltså

$$|\lambda - a_{ss}| \leq \sum_{\substack{k=1 \\ k \neq s}}^n |a_{sk}| |x_k| \leq \sum_{\substack{k=1 \\ k \neq s}}^n |a_{sk}| = R_s$$

Således ligger λ inom skivan C_s . Emedan λ är ett godtyckligt egenvärde, så måste alla egenvärden av A ligga innanför D .

¹Efter den vitryske matematikern *Semyon Aronovich Gershgorin* (1931)

Man kan också använda Gerschgorins cirklar för att studera effekten av störningar på elementen av A . Antag, att matrisen A har n linjärt oberoende egenvektorer, och betrakta den störda matrisen $A(\epsilon) = A + \epsilon B$ (B är en matris med samma norm som A). Vi kan uttrycka ekvationerna $Ax_i = \lambda_i x_i$, $i = 1, 2, \dots, n$ i formen $AX = X\Lambda$, $X = (x_1, x_2, \dots, x_n)$, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, där kolonnerna i X är egenvektorer för matrisen A . Eftersom egenvektorerna är linjärt oberoende, så existerar X^{-1} , och vi finner då att

$$X^{-1}A(\epsilon)X = X^{-1}AX + \epsilon X^{-1}BX = \Lambda + \epsilon C$$

har samma egenvärden som $A(\epsilon)$ ($A(\epsilon)$ och $X^{-1}A(\epsilon)X$ är likformiga, se sekt. 5.10). Genom att tillämpa Gerschgorins sats på denna matris får vi

$$|\lambda(\epsilon) - \lambda_i - \epsilon c_{ii}| \leq |\epsilon| \sum_{\substack{k=1 \\ k \neq i}}^n |c_{ik}|$$

för ett visst i , varav följer $|\lambda(\epsilon) - \lambda_i| \leq |\epsilon| \sum_{k=1}^n |c_{ik}|$ eller således $|\lambda_i(\epsilon) - \lambda_i| = \mathcal{O}(\epsilon)$. Om A är en symmetrisk matris, så kan X väljas ortogonal, och då är elementen i C av samma storleksordning som B :s element. Detta visar, att egenvärdesproblemet för en *symmetrisk* matris alltid är välartat.

Ibland är vi inte intresserade av alla egenvärden och egenvektorer av en matris. Vi skall nu beskriva en enkel metod att beräkna det till sitt absoluta värde största egenvärdet, jämte den tillhörande egenvektorn (**potensmetoden**).

Antag, att A är en $n \times n$ matris med n linjärt oberoende egenvektorer, och att egenvärdena uppfyller olikheterna $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$. En godtycklig vektor z_0 kan då uttryckas i formen $z_0 = \sum_{i=1}^n \alpha_i x_i$, där α_i är skalärer, som inte alla är lika med noll.

Om vi använder det iterativa schemat $z_k = Az_{k-1}$, $k = 1, 2, \dots$ där z_0 är en godtycklig vektor, så får vi $z_k = Az_{k-1} = A^2 z_{k-2} = \dots = A^k z_0 = \sum_{i=1}^n \alpha_i \lambda_i^k x_i$, genom att uttrycka z_0 med x_i som basvektorer.

Förutsatt att inte alla koefficienterna α_i är lika med noll, så kommer högra medlem av ovanstående ekvation småningom att domineras av termen $\alpha_1 \lambda_1^k x_1$, som innehåller det största egenvärdet. Om vi antar att $\alpha_1 \neq 0$, så kan vi följaktligen skriva

$$z_k = \lambda_1^k \left[\alpha_1 x_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k x_i \right] = \lambda_1^k (\alpha_1 x_1 + \epsilon_k),$$

om k är tillräckligt stort, varvid ϵ_k är en vektor med mycket små komponenter. Vektorn z_k är en god approximation till den onormerade egenvektorn x_1 endast om $\|\epsilon_k\|$ är tillräckligt liten.

Emedan $z_{k+1} = \lambda_1^{k+1}(\alpha_1 x_1 + \epsilon_{k+1})$ så gäller för varje i :

$$\frac{(z_{k+1})_i}{(z_k)_i} = \lambda_1 \left[\frac{\alpha_1 (x_1)_i + (\epsilon_{k+1})_i}{\alpha_1 (x_1)_i + (\epsilon_k)_i} \right] \rightarrow \lambda_1, \text{ om } k \rightarrow \infty$$

(observera, att $(z_k)_i$ betecknar den i :te komponenten av z_k).

I praktiken används följande räkneschema:

$$\begin{cases} y_k = Az_{k-1} \\ z_k = y_k / \|y_k\|_\infty \end{cases}$$

I detta fall gäller $z_k \rightarrow x_1 / \|x_1\|_\infty$ och $\|y_k\|_\infty \rightarrow \lambda_1$, då $k \rightarrow \infty$. Som en tillämpning skall vi beräkna det största egenvärdet för matrisen

$$A = \begin{pmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{pmatrix}.$$

Utgående från $z_0 = [001]^T$ får vi

k		z_k^T		$\ y_k\ _\infty$
0	0	0	1	1
1	0.5	1.0	0.25	4
2	0.5	1.0	0.8611	9
3	0.5	1.0	0.7306	11.4444
4	0.5	1.0	0.7335	10.9223
5	0.5	1.0	0.7493	11.0142
6	0.5	1.0	0.7501	10.9974
7	0.5	1.0	0.7500	11.0005
8	0.5	1.0	0.7500	10.9999

De exakta egenvärdena är i detta fall 11, -3 och -2 , vilket leder till små förhållanden $|\lambda_i/\lambda_1|$. Konvergensten är snabbast, då egenvärdena är mycket olika varandra. Konvergenstastigheten för det dominanta egenvärdet är beroende av förhållandet $|\lambda_2/\lambda_1|$.

Ett sätt att göra konvergensten snabbare är att tillämpa potensmetoden på matrisen $A - pI$ istället för A , där p är en parameter, som väljs på ett lämpligt sätt. Denna metod kallas **origoskift**, och kan vara mycket effektiv för vissa typer av matriser.

En matris av ordningen 4, som t.ex. har egenvärdena $\lambda_j = 15 - j$ har normalt en konvergensthastighet av storleksordningen $(13/14)^k$, men om vi väljer $p = 12$, blir konvergensthastigheten för matrisen $A - pI$ av storleksordningen $(\frac{1}{2})^k$. Man kan också använda sig av Aitkens accelerationsmetod.

I praktiken är metoden att bestämma dominanta egenvärden långsamma och rätt opålitliga. Det finns emellertid en ganska snabb iterativ metod, som mycket används på glesa matriser. Denna metod, som vi nu skall studera, kallas **invers iteration**.

Genom att ersätta A med A^{-1} i vårt iterationsschema ovan får vi en ny process som kan uttryckas i formen

$$\begin{cases} Ay^{(k+1)} = z^{(k)} \\ z^{(k+1)} = y^{(k+1)} / \|y^{(k+1)}\|_\infty \end{cases}$$

Den nya processen kommer således att konvergera mot det egenvärde av A , som har det *minsta* absoluta värdet.

För att visa detta skall vi anta att $z^{(0)} = \sum_{i=1}^n \alpha_i x_i$ där $Ax_i = \lambda_i x_i$. Då är

$$z^{(p)} = K_p \sum_{i=1}^n \alpha_i \lambda_i^{-p} x_i$$

där K_p är en skalningsfaktor.

Vektorn $z^{(p)}$ kommer alltså att domineras av den vektor x_n som svarar mot det minsta egenvärdet λ_n . Därför kommer räckan $z^{(p)}$ att konvergera mot en multipel av x_n då $p \rightarrow \infty$. För varje värde av j kommer dessutom att gälla att $z_j^{(p+1)} / z_j^{(p)} \rightarrow 1/\lambda_n$ då $p \rightarrow \infty$.

En allmännare (och mer praktisk) form av denna metod får man med origoskift, genom att utgå från matrisen $(A - qI)^{-1}$ istället för A^{-1} . Analogt med processen ovan får vi då

$$z^{(p)} = \sum_{i=1}^n \alpha'_i (\lambda_i - q)^{-p} x_i,$$

där $\alpha'_i = \alpha_i K_p$. Denna vektor kommer att domineras av den egenvektor x_d vars motsvarande egenvärde ligger närmast punkten q i det komplexa planet. Man inser lätt, att en relativt god approximation till en egenvektor leder till en snabb konvergens.

Man skulle tro att det uppstår svårigheter om begynnelsevektorn $z^{(0)}$ endast obetydligt beror av x_d , men så är inte fallet. Om vi t.ex. antar att $\lambda_d - q = \epsilon$, och att $|\lambda_i - q| > \theta$ då $i \neq d$, så får vi

$$z^{(p)} = \alpha'_d (\lambda_d - q)^{-p} x_d + \sum_{\substack{i=1 \\ i \neq d}}^n \alpha'_i (\lambda_i - q)^{-p} x_i = \frac{\alpha'_d}{\epsilon^p} x_d + \sum_{\substack{i=1 \\ i \neq d}}^n \alpha'_i (\lambda_i - q)^{-p} x_i$$

Härav följer

$$\frac{\epsilon^p}{\alpha'_d} z^{(p)} - x_d = \frac{\epsilon^p}{\alpha'_d} \sum_{\substack{i=1 \\ i \neq d}}^n \alpha'_i (\lambda_i - q)^{-p} x_i$$

eller alltså

$$\left\| \frac{\epsilon^p}{\alpha'_d} z^{(p)} - x_d \right\| \leq \frac{\epsilon^p}{|\alpha'_d|} \sum_{\substack{i=1 \\ i \neq d}}^n |\alpha'_i| \theta^{-p} = \left(\frac{\epsilon}{\theta} \right)^p \frac{1}{|\alpha'_d|} \sum_{\substack{i=1 \\ i \neq d}}^n |\alpha'_i|$$

Eftersom vi strävar efter konvergens mot λ_d så har q valts så, att λ_d är det egenvärde som ligger *närmast* q . Således måste $\epsilon < \theta$ gälla, och såvitt det inte finns flera egenvärden nära q , så är $\epsilon \ll \theta$. Således blir konvergensthastigheten proportionell mot ϵ/θ .

Metoden kan programmeras på följande sätt. De vektorer, som alstras vid invers iteration, är alla lösningar till ekvationssystemet $(A - qI)y^{(p+1)} = z^{(p)}$ för något visst värde av q . Observera, att matrisen $A - qI$ inte förändras. Om systemet löses genom LU -faktorisering behöver denna utföras endast en gång. Ekvationssystemet kommer då att reduceras till de triangulära systemen

$$\begin{cases} Lz = z^{(p)} \\ Uy^{(p+1)} = z, \end{cases}$$

där $LU = A - qI$.

Som vi har konstaterat, är valet av begynnelsevektor inte så viktigt. I standardmetoden brukar man faktiskt lämna bort den 0:te iterationen genom att sätta $Uy^{(1)} = e$, där $e = (1, 1, \dots, 1)^T$ (detta är ekvivalent med att välja $y^{(0)} = Le$).

5.10. Diagonaliseringsmetoder

Då man vill lösa det fullständiga egenvärdesproblemet för en matris, använder man sig vanligen av direkta metoder, som bygger på **likformighetstransformationer**. Om P är en godtycklig icke-singulär matris, så vet vi att både matrisen A och matrisen PAP^{-1} har samma egenvärden. I själva verket har de samma karaktäristiska ekvation, emedan

$$\det(\lambda I - PAP^{-1}) = \det(P(\lambda I - A)P^{-1}) = \det(P) \det(\lambda I - A) \det(P^{-1}) = \det(\lambda I - A).$$

Av $Ax = \lambda x$ följer dessutom att $(PAP^{-1})Px = \lambda(Px)$. Om x sålunda är en egenvektor till A , som kommer Px att vara en egenvektor till PAP^{-1} . Matriserna A och PAP^{-1} sägs vara *likformiga* och transformationen PAP^{-1} kallas för en likformighetstransformation. Om matrisen P är ortogonal, så påverkas inte egenvärdesproblemet's kondition därav.

För att diagonalisera matriser använder man vanligen en serie likformighetstransformationer som baserar sig på ortogonala matriser. En räcka matriser $A = A_0, A_1, A_2, \dots$ konstrueras då på följande sätt: $A_k = Q_k^T A_{k-1} Q_k$, där $Q_k^T Q_k = I$, $k = 1, 2, \dots$. Denna matrISRäcka bevarar en eventuell symmetri hos den ursprungliga matrisen, emedan $A_k^T = (Q_k^T A_{k-1} Q_k)^T = Q_k^T A_{k-1}^T Q_k$, $k = 1, 2, \dots$

De ortogonala transformationerna är vanligen av två slag: **plana rotationer**, eller **reflektioner**. En plan rotation i (p, q) -planet definieras genom matrisen $R_{p,q}(\phi)$, $\phi \leq \pi$, som skiljer sig från en enhetsmatris endast genom elementen $r_{pp} = r_{qq} = \cos \phi$, $r_{pq} = -r_{qp} = \sin \phi$. Ett exempel på en sådan matris har vi redan tidigare studerat.

Som vi ser, är $R_{p,q}^T(\phi) = R_{p,q}(-\phi)$ en matris av samma slag. Om A är en given matris och vi bildar produkten $A' = AR_{p,q}(\phi)$ så kommer endast elementen i *kolonnerna* p och q att förändras. Vi får $a'_{ij} = a_{ij}$ om $j \neq p, q$ samt

$$\begin{cases} a'_{ip} = \cos \phi a_{ip} - \sin \phi a_{iq} \\ a'_{iq} = \sin \phi a_{ip} + \cos \phi a_{iq} \end{cases}$$

Om man på samma sätt fullbordar likformighetstransformationen genom att bilda $A'' = R_{pq}^T A'$, så kommer endast elementen i *raderna* p och q att förändras.

I den andra transformationsmetoden, nämligen reflektionsmetoden, används ortogonala matriser av Householders typ (jfr. avsnitt 5.7): $P = I - 2ww^T$, där $w^T w = 1$. Som vi tidigare visat, är denna matris symmetrisk och ortogonal. Om man opererar med P på en matris $A = (a_1 a_2 \dots a_n)$, så ser man genom att bilda produkten $A' = PA$ att varje *kolonn* transformeras oberoende av de övriga enligt formeln $a'_k = (I - 2ww^T)a_k = a_k - 2(w^T a_k)w$. På samma sätt finner man genom att bilda produkten $A'' = A'P$ att raderna i detta fall transformeras oberoende av varandra.

Jacobis metod är den äldsta metoden att lösa det fullständiga egenvärdesproblemet². Den bygger på en räkka plana rotationer av typen $A_{k+1} = R^T(p, q)A_kR(p, q)$ där $R(p, q)$ är en rotation omfattande vinkeln ϕ i (p, q) -planet (jfr ovan). Elementen i A , som ligger ovanför huvuddiagonalen genomgås och det till sitt absoluta värde största elementet $a_{pq}^{(k)}$ söks upp. Rotationsvinkeln ϕ väljs därpå så, att elementet $a_{pq}^{(k+1)}$ blir *exakt noll*. De modifierade elementen beräknas sedan enligt formlerna

$$\left\{ \begin{array}{l} a_{ip}^{(k+1)} = a_{ip}^{(k)} \cos \phi + a_{iq}^{(k)} \sin \phi \\ a_{iq}^{(k+1)} = -a_{ip}^{(k)} \sin \phi + a_{iq}^{(k)} \cos \phi \\ a_{pp}^{(k+1)} = a_{pp}^{(k)} \cos^2 \phi + 2a_{pq}^{(k)} \cos \phi \sin \phi + a_{qq}^{(k)} \sin^2 \phi \\ a_{qq}^{(k+1)} = a_{pp}^{(k)} \sin^2 \phi - 2a_{pq}^{(k)} \cos \phi \sin \phi + a_{qq}^{(k)} \cos^2 \phi \\ a_{pq}^{(k+1)} = (a_{qq}^{(k)} - a_{pp}^{(k)}) \cos \phi \sin \phi + a_{pq}^{(k)} (\cos^2 \phi - \sin^2 \phi) \end{array} \right.$$

I formlerna gäller $i \neq p, q$; dessutom beaktas matrisens symmetri.

²C.G. Jacobi: *Über eine leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen*, Jour. f. reine u. angew. Math., Vol. XXX, 1846, 51-94

För att $a_{pq}^{(k+1)}$ skall försvinna, måste $-(a_{pp}^{(k)} - a_{qq}^{(k)}) \sin 2\phi + 2a_{pq}^{(k)} \cos 2\phi = 0$ gälla, dvs

$$\tan 2\phi = \frac{2a_{pq}^{(k)}}{a_{pp}^{(k)} - a_{qq}^{(k)}}$$

Vinkeln ϕ väljs så, att $-\frac{\pi}{4} \leq \phi \leq \frac{\pi}{4}$.

Av matriselementens transformationsformler följer, att matrisräckan A_k , $k = 1, 2, \dots$ faktiskt närmar sig en diagonal matris. Av definitionerna $E_k = \sum_{\substack{i,j=1 \\ i \neq j}}^n |a_{ij}^{(k)}|^2$ och $D_k = \sum_{i=1}^n |a_{ii}^{(k)}|^2$ följer nämligen

att $E_{k+1} = E_k - 2(a_{pq}^{(k)})^2$, som leder till $D_{k+1} = D_k + 2(a_{pq}^{(k)})^2$, vilket visar att matriserna närmar sig en diagonal matris. Man kan också enkelt visa, att matrisräckan närmar sig en *fixerad* diagonalmatris, dvs att $|a_{pp}^{(k+1)} - a_{pp}^{(k)}| \rightarrow 0$, då $k \rightarrow \infty$.

I sin ursprungliga form kan Jacobis metod vara något tidskrävande, pga att man måste uppsöka det största elementet utanför diagonalen. Därför väljer man vanligen någon annan strategi, t.ex. den cykliska metoden, där elementen genomgås i tur och ordning (oftast radvis), eller tröskelmetoden, där man nollställer alla element som överskrider ett visst tröskelvärde. Efter varje "svep" genom matrisen reduceras tröskelvärdet, t.ex. med hälften, varpå proceduren upprepas tills tröskelvärdet nått ett visst minimivärde.

Efter tillräckligt många rotationer får man en nästan diagonal matris. Vi har alltså $R^T A R = D$, där R är produkten av de plana rotationsmatriserna. Eftersom $A R = R D$, så kommer kolonnerna i R att innehålla egenvektorerna för A . Emedan alla transformationsmatriser är ortogonala, så kommer även R att vara ortogonal (även om A har närliggande egenvärden). Jacobis metod är mycket pålitlig, men inte särskilt snabb (det tar dubbelt så lång tid att räkna både egenvektorer och egenvärden som enbart egenvärden med denna metod).

De effektivaste metoderna nuförtiden för att diagonalisera matriser baserar sig på **Householders metod** för att överföra matrisen i tridiagonal form (om den är symmetrisk), eller Hessenbergs form (om den är osymmetrisk). En sådan matris är lättare att diagonalisera (t.ex. med QR -metoden). Bl.a. MATLAB-rutinen `eig` baserar sig på denna metod. Eftersom många av de matriser vi stöter på i fysiken är symmetriska, skall vi nöja oss med att beskriva hur en sådan kan transformeras till en tridiagonal form med Householders metod (1958).

I denna metod transformeras matrisen A enligt formeln $A_k = PA_{k-1}P$, $k = 1, 2, \dots$, där $P = I - 2w^{(k)}(w^{(k)})^T$ och vektorn $w^{(k)}$ är en enhetsvektor, vars komponenter ges av formeln $w^{(k)} = (0, 0, \dots, w_{k+1}^{(k)}, \dots, w_n^{(k)})^T$. Matrisen P kommer att transformera A -matrisens element i raderna $k + 1$ t.o.m. n och i kolonnerna $k + 1$ t.o.m. n .

Vektorn $w^{(k)}$ beräknas på följande sätt. Antag, att x är den k :te kolonnen i A_k . Vi väljer $w^{(k)}$ så, att elementen med radnummer $k + 2$ t.o.m. n i den k :te kolonnen av A_{k+1} är lika med noll. Vektorn uppdelas sedan på följande sätt: $x^T = (x'^T, x_{k+1}, y^T)$, där x'^T är en radvektor med k element och y^T en radvektor med $n - (k + 1)$ element. På motsvarande sätt uppdelas w : $w^{(k)T} = (0, \dots, w_{k+1}^{(k)}, u^T)$. För att kunna beräkna w räcker det att studera vektorn Px , som har formen

$$\begin{pmatrix} x' \\ x_{k+1} \\ y \end{pmatrix} - 2 \begin{pmatrix} \mathbf{0} \\ w_{k+1}^{(k)} \\ u \end{pmatrix} (w_{k+1}^{(k)} x_{k+1} + u^T y) = \begin{pmatrix} x' \\ x_{k+1} - 2\alpha w_{k+1}^{(k)} \\ y - 2\alpha u \end{pmatrix},$$

där $\alpha = (w_{k+1}^{(k)} x_{k+1} + u^T y)$. För att elementen med index $k + 2$ t.o.m. n i Px skall försvinna, måste man uppenbarligen välja $u = \gamma y$, där γ är en konstant, och sätta $1 - 2\alpha\gamma = 0$. Vi fordrar också att w skall vara en enhetsvektor, vilket leder till villkoret $|\gamma|^2 y^T y + |w_{k+1}^{(k)}|^2 = 1$.

Om vi väljer $w_{k+1}^{(k)} = \gamma v_{k+1}^{(k)}$, så följer av villkoren ovan

$$\begin{cases} 1 - \gamma^2(2v_{k+1}^{(k)}x_{k+1} + 2y^T y) = 0 \\ 1 - \gamma^2[(v_{k+1}^{(k)})^2 + y^T y] = 0 \end{cases}$$

varur γ kan elimineras.

Härav följer $v_{k+1}^{(k)} = x_{k+1} \pm \sqrt{x_{k+1}^2 + y^T y} = x_{k+1} \pm \beta$, där förtecknet bör väljas i överensstämmelse med x_{k+1} . Faktorn γ bestäms slutligen ur ekvationen $\gamma = [2\beta(\beta \pm x_{k+1})]^{-1/2}$ varur $w_{k+1}^{(k)}$ kan beräknas. Processen kan arrangeras så, att endast en kvadratrots behövs beräknas vid varje transformation.

För att beräkna egenvärdena av den tridiagonala matris, som uppstår efter Householder-transformationen, användes tidigare en metod, där egenvärdenas fördelning inom olika intervall studeras med hjälp av **Sturms bisektionsmetod**. Numera använder man oftast den nedan beskrivna QR -metoden (Francis,1961), som är effektivare.

Genom upprepade elementära rotationer kan en godtycklig symmetrisk matris förvandlas till en övre triangulär matris, m.a.o. vi kan uppdelas matrisen A i formen $A_1 = Q_1 R_1$, och därpå konstruera en ny matris $A_2 = R_1 Q_1$, som kan erhållas genom en likformighetstransformation $Q_1^T A_1 Q_1$ ur A_1 (visas genom substitution). Vi får alltså en matrISRäcka $A_k = Q_k R_k$, $A_{k+1} = R_k Q_k = Q_k^T A_k Q_k, \dots$

I praktiken beräknas den ortogonala matrisen Q_k^T så att $Q_k^T A_k = R_k$.

Eftersom $A_{k+1} = R_k Q_k = Q_k^T A_k Q_k$, så får vi genom att upprepa transformationerna

$$A_{k+1} = Q_k^T A_k Q_k = Q_k^T Q_{k-1}^T A_{k-1} Q_{k-1} Q_k = \dots = (Q_k^T \dots Q_1^T) A_1 (Q_1 \dots Q_k)$$

eller alltså $Q_1 \dots Q_k A_{k+1} = A_1 Q_1 \dots Q_k$. Man kan också visa, att om A_1 speciellt är en tridiagonal matris, så kommer alla matriserna A_k att vara tridiagonala. Detta leder till att QR -algoritmen blir speciellt effektiv för tridiagonala matriser.

Man kan visa, att matrisräckan A_k , $k = 1, 2, \dots$ kommer att konvergera mot en övre triangulär matris, där egenvärdena befinner sig på diagonalen. Om förhållandet mellan några av egenvärdena $|\lambda_{i+1}/\lambda_i|$ är nära 1, så blir konvergensen långsam (liksom i den iterativa potensmetoden), och därför brukar man kombinera QR -metoden med origoskift: $A_k - p_k I = Q_k R_k$ och $A_{k+1} = p_k I + R_k Q_k$, där p_k är en skiftparameter, som väljs så att konvergensen blir snabbare.