ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Σχολή Μηχανολόγων Μηχανικών
*Εργαστήριο Θερμικών Στροβιλομηχανών*
*Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής & Βελτιστοποίησης*

# Η (Γραμμική) Μέθοδος GMRES

**Κυριάκος Χ. Γιαννάκογλου**
**Καθηγητής ΕΜΠ**

**kcgcours@central.ntua.gr**

**http://velos0.ltt.mech.ntua.gr/research/**

# Γενικά – το Concept

GMRES: Generalized Minimal RESidual Technique

(Youcef Saad, Yale Univ)

a Krylov subspace method

Concept: $Ax = b$, αρχικοποίηση $x^0$

προσεγγιστική λύση $x = x^0 + \sum_{i=1}^{\dots} \beta_i v^i$

οπου: (α) $v^1, v^2, \dots$ μια ορθοκανονική βάση $\begin{cases} (v^i, v^j) = 0, & i \neq j \\ (v^i, v^j) = 1, & i = j \end{cases}$

(β) $\beta_1, \beta_2, \dots$ συντελεστές πού $\|b - Ax\| \to \min.$

# Δημιουργία Βάσης – Ο Αλγόριθμος Arnoldi

| ΑΛΓΟΡΙΘΜΟΣ (Α1) | ΑΛΓΟΡΙΘΜΟΣ (Α2) |
|---|---|
| $v^1 = \dfrac{r^0}{\lVert r^0 \rVert}$ | $v^1 = \dfrac{r^0}{\lVert r^0 \rVert}$ |
| **DO j=1,m** | **DO j=1,m** |
| $w^j = Av^j$ | $w^j = Av^j$ |
| **do i=1,j** | **do i=1,j** |
| $h_{ij} = (w^j, v^i)$ | $h_{ij} = (w^j, v^i)$ |
| **enddo** | $w^j = w^j - h_{ij} v^i$ ← |
| $w^j = w^j - \sum_{i=1}^{j} h_{ij} v^i$ ← | **enddo** |
| $h_{j+1,j} = \lVert w^j \rVert_2$ | $h_{j+1,j} = \lVert w^j \rVert_2$ |
| $v^{j+1} = \dfrac{w^j}{h_{j+1,j}}$ | $v^{j+1} = \dfrac{w^j}{h_{j+1,j}}$ |
| **ENDDO** | **ENDDO** |

# Δημιουργία Βάσης – Ο Αλγόριθμος Arnoldi

Γιά βάση $m$ :

$$\underbrace{v^1, \dots, v^{m+1}}_{\text{Διανύσματα βάσης}}$$

Διανύσματα βάσης

( σύν ένα !!!)

$$\bar{H}_m = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,m} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,m} \\ & h_{3,2} & & h_{3,m} \\ & & & \vdots \\ & & & h_{m,m} \\ & & & h_{m+1,m} \end{bmatrix} \qquad \begin{matrix} (H_m) \\ \uparrow \end{matrix}$$

$(m+1 * m)$ Hessemberg matrix

# Δημιουργία Βάσης – Ο Αλγόριθμος Arnoldi

Μηδενικό Βήμα: $r^0 = b - Ax^0 \quad , \quad v^1 = \dfrac{r^0}{\|r^0\|}$

Πρώτο Βήμα:

$$h_{1,1} = (Av^1, v^1)$$

$$\hat{v}^2 = Av^1 - h_{1,1}v^1$$

$$h_{1,2} = \|\hat{v}^2\|$$

$$v^2 = \frac{\hat{v}^2}{h_{1,2}}$$

(m )-ιοστό Βήμα:

$$h_{1,m} = (Av^m, v^1) \; , \; h_{2,m} = (Av^m, v^2) \; , \; \dots , \; h_{m,m} = (Av^m, v^m)$$

$$\hat{v}^{m+1} = Av^m - \sum_{i=1}^{m} h_{i,m} v^i \longrightarrow \text{προσωρινό}$$

$$h_{m+1,m} = \|\hat{v}^{m+1}\|$$

$$v^{m+1} = \frac{\hat{v}^{m+1}}{h_{m+1,m}}$$

# Δημιουργία Βάσης – Ο Αλγόριθμος Arnoldi

"Επίδειξη" Ορθοκανονικότητας:

$$0=(v^2,v^1)=(Av^1-h_{1,1}v^1,v^1)=(Av^1,v^1)-h_{1,1}(v^1,v^1)\Rightarrow h_{1,1}=(Av^1,v^1)$$

$$0=(v^3,v^1)=(Av^2-h_{1,2}v^1-h_{2,2}v^2,v^1)=(Av^2,v^1)-h_{1,2}(v^1,v^1)-h_{2,2}(v^2,v^1)\Rightarrow$$

$$\Rightarrow h_{1,2}=(Av^2,v^1)$$

Σύνοψη:

- Μηδενικό βήμα + άλλα m βήματα →

  $v^1, v^2, v^3, \ldots, v^m, v^{m+1}$

- Υπολογιστικό Κόστος?

  Matrix-Vector Multiplications! Πόσες?

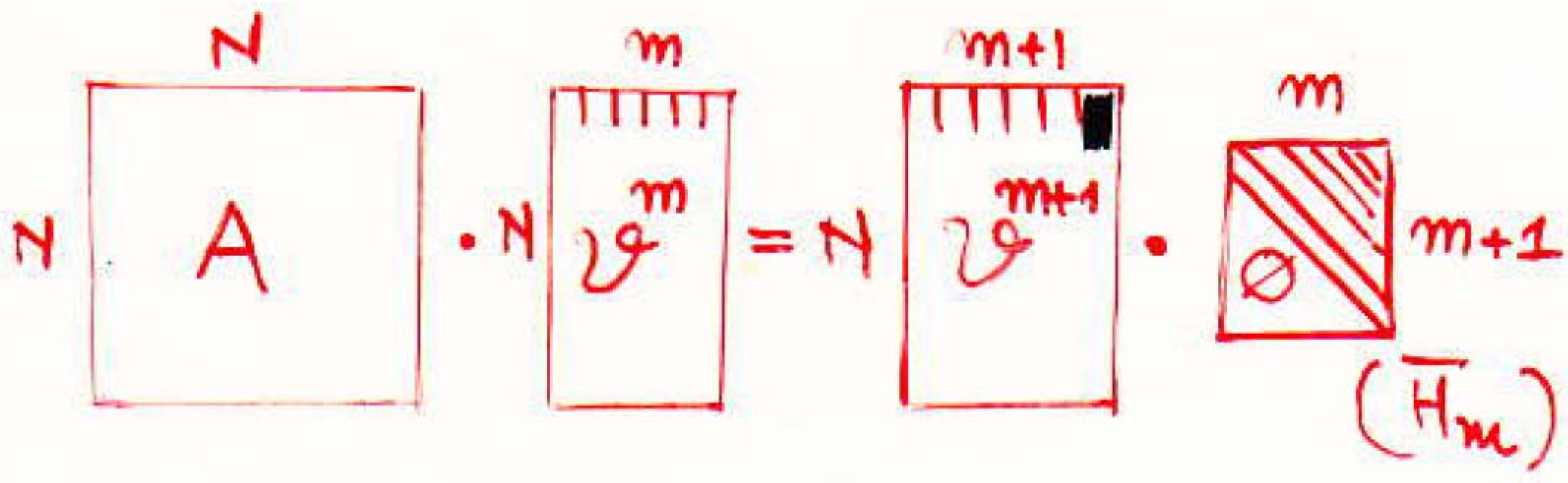
# Βασικές Εξισώσεις από τον Αλγόριθμο Arnoldi

$$A\vartheta^m = \vartheta^m H_m + w^m e_m^T \qquad (1)$$

$$= \vartheta^{m+1}\bar{H}_m \qquad (2)$$

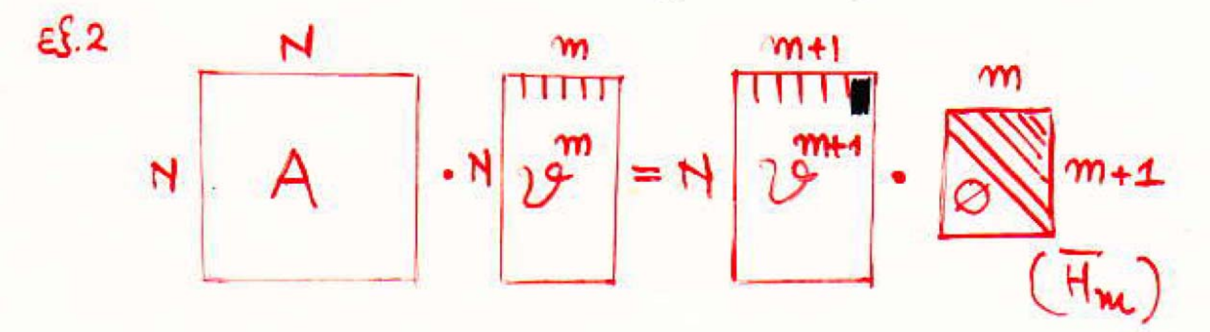
$$(\vartheta^m)^T A\, \vartheta^m = H_m \qquad (3)$$

εξ.2



# Απόδειξη Εξίσωσης (2)

$$\boxed{A\vartheta^{m} = \vartheta^{m+1}\,\overline{H}_m}$$




Στόν αλγόριθμο Arnoldi είχαμε ότι :

$$\left.\begin{aligned} w^j &= Av^j - \sum_{i=1}^{j} h_{i,j} v^i \\ v^{j+1} &= \frac{w^j}{h_{j+1,j}} \end{aligned}\right\} \Rightarrow Av^j = h_{j+1,j}\, v^{j+1} + \sum_{i=1}^{j} h_{i,j} v^i$$

δηλαδή, συνοψίζοντας σέ μία αθροιση : $Av^j = \sum_{i=1}^{j+1} h_{i,j} v^i$

Σέ μητρική γραφή είναι ή πρός απόδειξη εξίσωση (2) , $A\vartheta^{m} = \vartheta^{m+1}\,\overline{H}_m$ .

# Απόδειξη Εξίσωσης (1)

$$A v^m = v^m H_m + w^m e_m^T \qquad (1)$$

$$= v^{m+1} \bar{H}_m \qquad (2)$$

$$(v^m)^T A v^m = H_m \qquad (3)$$

Γνωρίζοντας ότι η τελευταία γραμμή του $\bar{H}_m$ περιέχει μόνο το στοιχείο $h_{m+1,m}$ αυτή η σχέση ξαναγράφεται ως

$$A v^m = \underbrace{v^m H_m + h_{m+1,m} v^{m+1}}_{} = v^m H_m + w^m e_m^T \; , \quad (1) \quad \text{ο.ε.δ.}$$

# Απόδειξη Εξίσωσης (3)

$$A \vartheta^m = \vartheta^m H_m + w^m e_m^T \qquad (1)$$

$$= \vartheta^{m+1} \bar{H}_m \qquad (2)$$

$$(\vartheta^m)^T A \vartheta^m = H_m \qquad (3)$$

Τέλος, πολλαπλασιάζουμε την (1) με το $(\vartheta^m)^T$ και έχουμε

$$(\vartheta^m)^T A \vartheta^m = \underbrace{(\vartheta^m)^T \vartheta^m}_{\substack{\text{Ταυτοτικός} \\ \text{Πίνακας}}} H_m + \underbrace{(\vartheta^m)^T w^m e_m^T}_{\substack{\text{Μηδέν, λόγω ορθογωνιότητας} \\ \text{των } (v^1, v^2, \dots, v^m)}} = H_m \text{, (3) ο.ε.δ.}$$

# Restarted GMRES

ΒΗΜΑ 1 : Υπολογισμός $r^0 = b - Ax^0$ , $v^1 = \frac{r^0}{\|r^0\|_2}$

ΒΗΜΑ 2 : Υπολογισμός m-1 (ουσιαστικά όμως m) διανυσμάτων βάσης $v^2, v^3, \dots, v^m$ (ουσιαστικά όμως και του $v^{m+1}$) ως εξής:

FOR $j = 1, 2, \dots, m$

υπολογισμός του $w^j = Av^j$

FOR $i = 1, 2, \dots, j$

$$h_{i,j} = (w^j, v^i)$$

END

$$v^{j+1} = w^j - \sum_{i=1}^{j} h_{i,j} v^i$$

$$h_{j+1,j} = \| v^{j+1} \|_2$$

Κανονικοποίηση του $v^{j+1}$ : $v^{j+1} = \frac{v^{j+1}}{h_{j+1,j}}$

END

ΒΗΜΑ 3 : Υπολογισμός, μέσω επίλυσης ενός "μικρού" προβλήματος ελαχιστοποίησης, των m συντελεστών $\beta_1, \dots, \beta_m$ και δημιουργία της προσεγγιστικής λύσης

$$x^m = x^0 + \sum_{i=1}^{m} \beta_i v^i$$

# Υπολογισμός των συντελεστών $\beta_i$

"Κτίζοντας" τη νέα προσεγγιστική λύση ως $x^m = x^0 + \sum_{i=1}^{m} \beta_i v^i$, θα απαιτήσουμε το υπόλοιπο $r^m$ του $x^m$ να έχει ελάχιστη τιμή νόρμας-2.

Απαιτούμε λοιπόν

$$\min \| A x^m - b \|_2 = \min \| A x^0 + \sum_{i=1}^{m} \beta_i A v^i - b \|_2 = \min \| r^0 - \sum_{i=1}^{m} \beta_i A v^i \|_2$$

$$\boxed{\min \| r^0 - (A \vartheta^m) B_m \|_2}$$

$\beta_i$, $i=1,m$ είναι τα στοιχεία ενός διανύσματος στήλης (διάστασης $m \times 1$) $B_m$

$$A \vartheta^m = \vartheta^{m+1} \overline{H}_m \Rightarrow \boxed{\min \| r^0 - \vartheta^{m+1} \overline{H}_m B_m \|_2}$$

# Υπολογισμός των συντελεστών $\beta_i$

$$\boxed{\min \left\| r^0 - \vartheta^{m+1} \overline{H}_m B_m \right\|_2}$$

$$r^0 = \|r^0\|_2 \, v^1$$

$$v^1 = \vartheta^{m+1} e_1$$

$$e_1 = \begin{array}{cc} \begin{bmatrix} 1 \\ \emptyset \\ \emptyset \\ \vdots \\ \emptyset \end{bmatrix} & m+1 \end{array}$$

$$\min \left\| \, \|r^0\|_2 \vartheta^{m+1} e_1 - \vartheta^{m+1} \overline{H}_m B_m \right\|_2$$

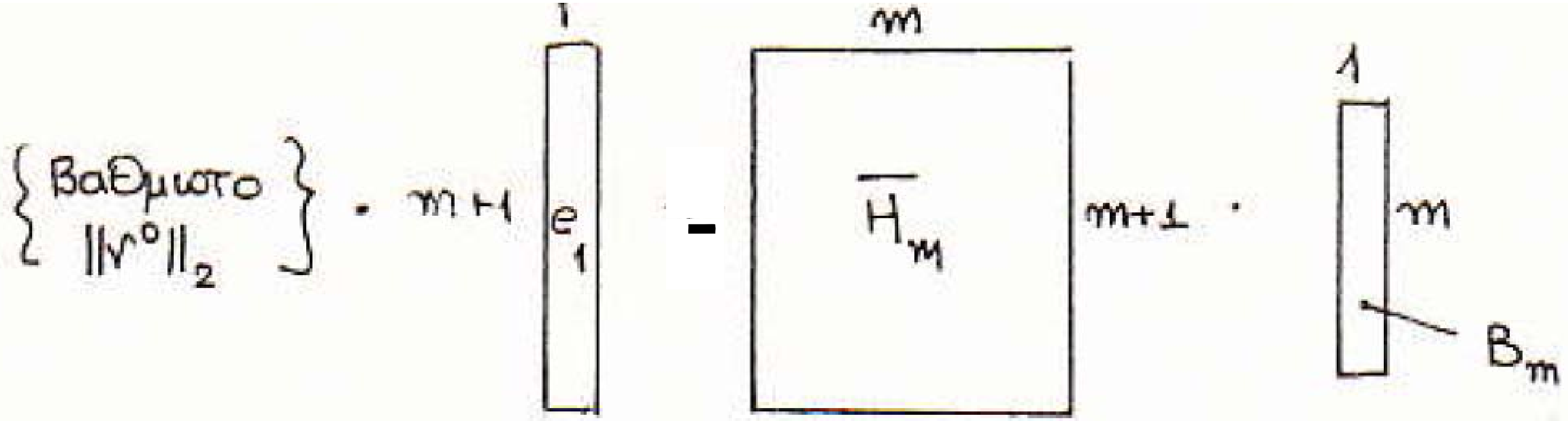
$$= \min \left\| \vartheta^{m+1} \left\{ \|r^0\|_2 e_1 - \overline{H}_m B_m \right\} \right\|_2$$

# Υπολογισμός των συντελεστών $\beta_i$

$$\min \left\| \vartheta^{m+1} \left\{ \|r^0\|_2 e_1 - \bar{H}_m B_m \right\} \right\|_2$$

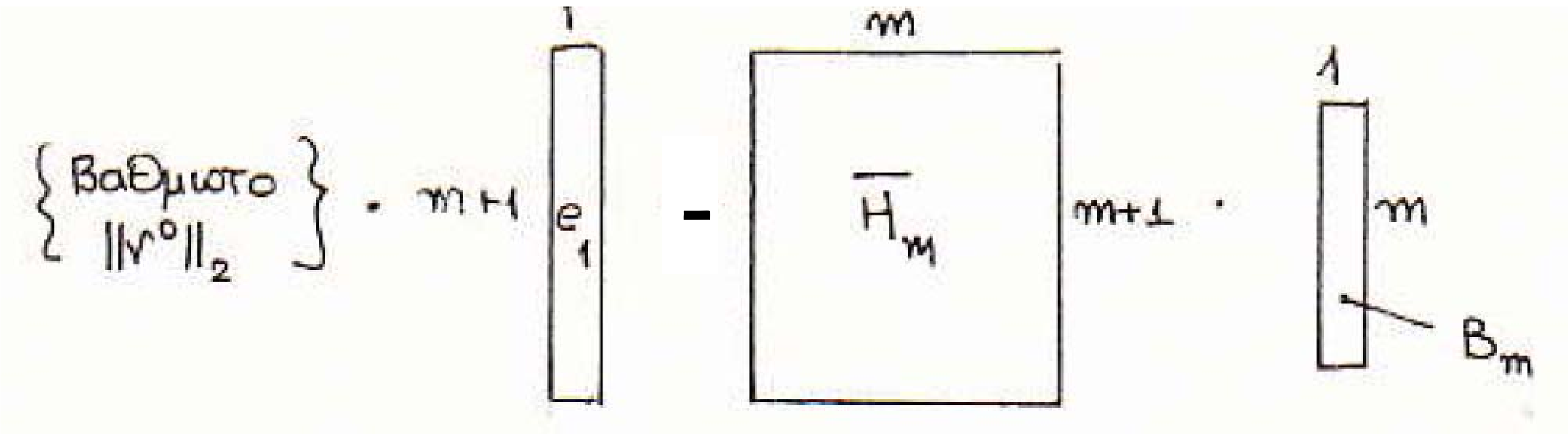
Επειδή τα διανύσματα στήλης πού συνθέτουν τό $\vartheta^{m+1}$ είναι ανεξάρτητα των $\beta_i$, απαιτούμε ουσιαστικά τη συνθήκη:

$$\min \left\| \, \|r^0\|_2 e_1 - \bar{H}_m B_m \right\|_2$$



# Υπολογισμός των συντελεστών $\beta_i$

$$\boxed{\min \left\| \, \|v^0\|_2 \, e_1 - \bar{H}_m B_m \right\|_2}$$




Πρόκειται γιά (m+1) εξισώσεις πού πρέπει νά ικανοποιηθούν έχοντας γιά αγνώστους (m) (μόνο!) ποσότητες $(\beta_1, \cdots, \beta_m)$. Τό πρόβλημα αντιμετωπίζεται μέ τήν τεχνική τών ελαχίστων τετραγώνων (Least Squares Problem)

<u>Πρακτικά:</u>

Θεωρείστε μητρώα στροφής $F_i$, όπως λ.χ. τα παρακάτω:

$$F_1 = \begin{array}{|c|c|c|c|l} c_1 & -s_1 & 0 & 0 & \\ \hline s_1 & c_1 & 0 & 0 & \\ \hline 0 & 0 & 1 & 0 & \cdots \\ \hline 0 & 0 & 0 & 1 & \\ \hline & & & & \end{array} \quad , \quad F_2 = \begin{array}{|c|c|c|c|l} 1 & 0 & 0 & 0 & \\ \hline 0 & c_2 & -s_2 & 0 & \\ \hline 0 & s_2 & c_2 & 0 & \cdots \\ \hline 0 & 0 & 0 & 0 & \\ \hline & & & & \end{array}$$

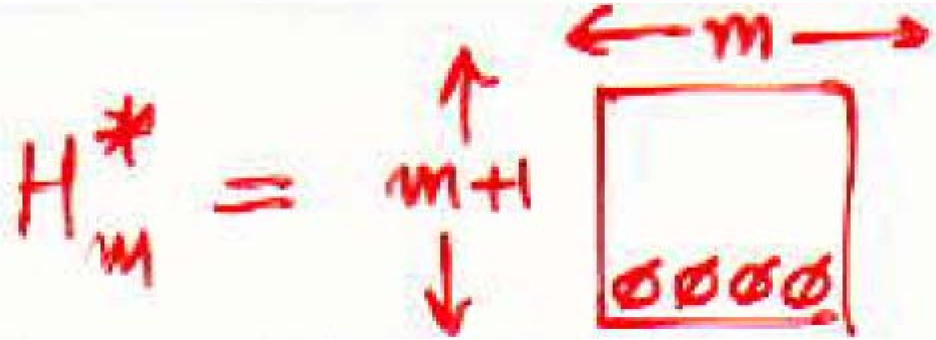
Εφαρμόζοντας λ.χ. το $F_1$ στο $\bar{H}_m$ μπορεί να απαλειφθεί το $h_{2,1}$ (δηλ. το κάτω της διαγωνίου στοιχείο της πρώτης στήλης), αφού τα $c_1, s_1$ να υπολογισθούν κατάλληλα (πως; με τη στροφή $F_1$ ποια άλλα στοιχεία του $\bar{H}_m$ αλλάζουν;).

Μέ m εν σειρά εφαρμογές τών τελεστών στροφής $F_1, F_2, \ldots, F_m$ στό μητρώο $\overline{H}_m$ μπορεί να προκύψει ένας (m×m) άνω τριγωνικός πίνακας $H_m^*$. (είναι (m+1×m) αλλά η τελευταία γραμμή είναι "όλο μηδενικά"). Συμβολικά:

$$(F_m \cdot F_{m-1} \cdots F_1)\,\overline{H}_m = \overline{H}_m^*$$

Οι m διαδοχικές στροφές μπορούν να συναθροιστούν στην τελεστή Q και είναι:

$$Q\,\overline{H}_m = \overline{H}_m^*$$




όπου δείχνεται εύκολα ότι ισχύει : $Q^H Q = I$

($Q^H$ = transpose conjugate : $Q^H = \overline{Q}^T = \overline{Q^T}$ )

# Υπολογισμός των συντελεστών $\beta_i$

Η ελαχιστοποίηση διατυπώνεται τότε ως $\min \left\| Q \, \|r^0\|_2 e_1 - Q \bar{H}_m B_m \right\|_2$

$$\min \left\| \bar{g}_m - \bar{H}_m^* B_m \right\|_2 \qquad (2)$$

όπου

$$\bar{g}_m = Q \, \|r^0\|_2 e_1 = (\gamma_1, \gamma_2, \cdots, \gamma_{m+1})^T$$

Η ελαχιστοποίηση πού εκφράζει ή σχέση (2) περιορίζεται στην επίλυση ενός $(m \times m)$ γραμμικού προβλήματος με αγνωστους τά $\beta_1, \beta_2, \ldots, \beta_m$.

Απλά λοιπόν λύνουμε το σύστημα

$$\boxed{H_m^* B_m = g_m}$$

# Υπολογισμός των συντελεστών $\beta_i$

$$\boxed{H_m^* B_m = g_m}$$

όπου ας θεωρήσουμε ότι τό $H_m^*$ είναι ένα $(m \times m)$ μητρώο που προκύπτει από τό $\overline{H}_m^*$ απαλείφοντας τη (μηδενική) τελευταία γραμμή του, ενώ

$$g_m = (\gamma_1, \gamma_2, \cdots, \gamma_m)^T$$

# (Σημαντική) Πρόταση:

Ισχύει (καί είναι πρακτικά ιδιαίτερα χρήσιμο όταν προγραμματίζουμε GMRES)

ότι:

$$\gamma_{m+1} = \left\| \, \|r^0\|_2 e_1 + \overline{H}_m B_m \right\|_2$$

Απόδειξη:

Μέ τόν υπολογισμό τής προσεγγιστικής λύσης $x^m$, το αντίστοιχο υπόλοιπο γίνεται:

$$r^m = b - Ax^m = b - A(x^0 + \vartheta^m B_m) = r^0 - A\vartheta^m B_m =$$

$$= r^0 - \vartheta^{m+1}\overline{H}_m B_m = \|r^0\|_2 \vartheta^{m+1} e_1 - \vartheta^{m+1}\overline{H}_m B_m =$$

$$= \vartheta^{m+1}\left\{ \|r^0\|_2 e_1 - \overline{H}_m B_m \right\}$$

# (Σημαντική) Πρόταση:

$$r^m = \vartheta^{m+1} \{ \|r^0\|_2 e_1 - \overline{H}_m B_m \}$$

Επειδή $Q^T Q = I$ γράφουμε:

$$r^m = \vartheta^{m+1} Q^T Q \{ \|r^0\|_2 e_1 - \overline{H}_m B_m \} =$$

$$= \vartheta^{m+1} Q^T \{ \overline{g}_m - \overline{H}^*_m B_m \}$$

Στο πρόβλημα ελαχιστοποίησης επέλεξα να λύσω την $H^*_m B_m = g_m$ πού πρακτικά απαλείφει κάθε συνιστώσα του $\overline{g}_m = (\gamma_1, \gamma_2, \cdots, \gamma_{m+1})^T$ εκτός από την τελευταία (τη $\gamma_{m+1}$). Άρα

$$r^m = b - Ax^m = \vartheta^{m+1} Q^T (\gamma_{m+1} e_{m+1})$$

# (Σημαντική) Πρόταση:

$$r^m = b - Ax^m = V^{m+1} Q^T \left( \gamma_{m+1} e_{m+1} \right)$$

Εκμεταλλευομενοι ομως οτι η βαση που δημιουργησαμε είναι ορθοκανονική

Προκύπτει εύκολα ότι $V^{m+1} Q^T = I$ και έτσι

$$\underline{\|r^m\|_2} = \| b - Ax^m \|_2 = \underline{\gamma_{m+1}} \quad , \text{ο.ε.δ.}$$

περιέχει χρήσιμη πληροφορία –

# Preconditioned Restarted GMRES

Ένα M είναι ο προδιαθέτης, οι αλλαγές που απαιτούνται στον αλγόριθμο που προηγουμένα δόθηκε για τη Restarted GMRES είναι οι εξής:

(α) στο <u>ΒΗΜΑ 1</u>: πρέπει να υπολογισθεί το υπόλοιπο της προδιάθεσης (preconditioned residual) ως

$$r^0 = M^{-1}(b - Ax^0) \quad , \text{ή αλλοιώς} \quad M\, r^0 = b - Ax^0$$

και επίσης να κανονικοποιηθεί ως $$v^1 = \frac{r^0}{\|r^0\|_2}$$

(β) στο <u>ΒΗΜΑ 2</u>: να χρησιμοποιείται $(\forall j)$ η ποσότητα

$$w^j = M^{-1} A v^j$$

πού ουσιαστικά απαιτεί τη "λύση" του συστήματος

$$M w^j = A v^j$$

# Preconditioned Restarted GMRES

Παρατηρηση : Το πόσο επιπλέον "στοιχίζουν" οι πράξεις μετά μητρωο M καθορίζουν και το αν "αξίζει" ή οχι να χρησιμοποιείται ο προδιαθέτης που διαλέξαμε.

π.χ. $M = LU$

λύνεται το $LUw^j = (Av^j) =$ γνωστο·

με δυο βήματα

1) $L\varphi = Av^j \longrightarrow \varphi$

$Uw^j = \varphi \longrightarrow w^j$

# Arnoldi-Full Orthogonalization Method (FOM)

ΒΗΜΑ 1 : Αρχική επιλογή $x^0$. Εύρεση αρχικού υπολοίπου $r^0 = b - Ax^0$

Υπολογισμός $v^1 = r^0 / \|r^0\|_2$

Μηδενισμός του μητρώου (διάστασης $m \times m$) $H_m$ , $h_{i,j}=0$

ΒΗΜΑ 2 :

FOR $j=1,2,\dots,m$

υπολογισμός του $w^j = Av^j$

FOR $i=1,2,\dots,j-1$

$h_{i,j} = (w^j, v^i)$

$w^j = w^j - h_{i,j} v^i$

END

υπολογισμός του $h_{j+1,j} = \|w^j\|_2$

Αν $h_{j+1,j}=0$ , εκτελείται το ΒΗΜΑ 3

υπολογισμός του $v^{j+1} = \dfrac{w^j}{h_{j+1,j}}$

END

μητρωική γραφή

ΒΗΜΑ 3 : Σχηματισμός της προσεγγιστικής λύσης $x^m = x^0 + \vartheta^m B_m$

(όπου $B_m$ είναι το διάνυσμα στήλης με τους m συντελεστές $\beta_j$)

Ισχύει: $B_m = H_m^{-1} \left( \|r^0\|_2 e_1 \right)$

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$