μετασχηματισμό τέτοιο ώστε επιδρώντας στο λάθος πρόβλεψης $\mathbf{e}$, $(\mathbf{e}^{'} = \mathbf{Te})$ να οδηγεί σε ελαχιστοποίηση του $E^{'} = \mathbf{e}^{'T}\mathbf{e}^{'}$ όταν ελαχιστοποιείται και το *E,* και να μετασχηματίζει τον πίνακα **G** στον πίνακα $\mathbf{G}'$ που να έχει άνω τριγωνική μορφή, θα έχομε

$$\mathbf{e}^{'} = \mathbf{Te} = \mathbf{T}(\mathbf{d} - \mathbf{Gm}) = \mathbf{Td} - \mathbf{TGm} = \mathbf{d}' - \mathbf{G}^{'}\mathbf{m} \qquad (6.11)$$

που αναλυτικά γράφεται ως

$$\begin{bmatrix} e_1^{'} \\ e_2^{'} \\ e_3^{'} \\ . \\ e_M^{'} \\ . \\ . \\ e_N^{'} \end{bmatrix} = -\begin{bmatrix} G_{11}^{'} & G_{12}^{'} & G_{13}^{'} & G_{14}^{'} & . & G_{1M}^{'} \\ 0 & G_{22}^{'} & G_{23}^{'} & G_{24}^{'} & . & G_{2M}^{'} \\ 0 & 0 & G_{33}^{'} & G_{34}^{'} & . & G_{3M}^{'} \\ . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & G_{MM}^{'} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} m_1^{'} \\ m_2^{'} \\ m_3^{'} \\ . \\ m_M^{'} \end{bmatrix} + \begin{bmatrix} d_1^{'} \\ d_2^{'} \\ d_3^{'} \\ . \\ d_M^{'} \\ . \\ . \\ d_N^{'} \end{bmatrix} \qquad (6.12)$$

Στην περίπτωση αυτή βλέπομε ότι οποιαδήποτε επιλογή των $m_i^{'est}$ δεν επηρεάζει τις τελευταίες *N-M* εξισώσεις . Επομένως μπορούμε να ζητήσομε να μηδενίζονται οι πρώτες *M* τιμές του $\mathbf{e}'$, οπότε ικανοποιούνται ακριβώς οι πρώτες Μ εξισώσεις του συστήματος που παίρνουν τη μορφή $\mathbf{d}' = \mathbf{G}^{'}\mathbf{m}^{'}$. Επί πλέον λόγω της μορφής του πίνακα $\mathbf{G}'$ μπορούμε να λύσουμε το σύστημα με αντίστροφη διαδικασία Το συνολικό λάθος λοιπόν είναι :

$$E' = \sum_{i=M+1}^{N} e_i^{'2} = \sum_{i=M+1}^{N} d_i^{'2} . \qquad (6.13)$$

Με το μετασχηματισμό, διαχωρίσαμε το αντίστροφο πρόβλημα σε ένα μέρος το οποίο αφορά μετρήσεις που ικανοποιούνται επακριβώς και σε ένα άλλο μέρος που αφορά μετρήσεις που δεν μπορούν να ικανοποιηθούν. Η λύση επιλέγεται έτσι ώστε να ελαχιστοποιείται το λάθος πρόβλεψης και συνεπώς είναι λύση ελαχίστων τετραγώνων.

6.2.3 Το μεικτά ορισμένο πρόβλημα.

Σε ένα μεικτά ορισμένο πρόβλημα της μορφής **Gm=d** κάποιοι γραμμικοίΙ συνδυασμοί παραμέτρων είναι υπερ-ορισμένοι και κάποιοι υπο-ορισμένοι.

Εάν ένα πρόβλημα είναι σε κάποιο βαθμό υποορισμένο, η εξίσωση **Gm=d** περιέχει πληροφορία για μερικές μόνο από τις παραμέτρους. Οι παράμετροι αυτές θεωρούμε ότι περιλαμβάνονται σε ένα υπόχωρο $S_p(\mathbf{m})$ του χώρου των παραμέτρων. Για το συμπλήρωμα του υπόχωρου αυτού ως προς το χώρο των παραμέτρων, ο πίνακας **G** δεν παρέχει καμμία πληροφορία. Ο υπόχωρος αυτός χαρακτηρίζεται ως $S_0(\mathbf{m})$.

Αντίστοιχα σε ένα υπερορισμένο πρόβλημα το γινόμενο **Gm** μπορεί να μην μπορεί να παράγει το χώρο $S(\mathbf{d})$ ανεξά (6.15ρτητα από την επιλογή των παραμέτρων. Το πολύ να μπορεί να ικανοποιηθεί ένα μέρος των μετρήσεων που χαρακτηρίζεται ως $S_p(\mathbf{d})$ σε αντίθεση με τον υπόχωρο $S_0(\mathbf{d})$ που περιλαμβάνει τις μετρήσεις που δεν μπορούν να ικανοποιηθούν.

Θεωρώντας τώρα παραμέτρους και μετρήσεις που ανήκουν στους αντίστοιχους υπόχωρους με τους ανωτέρω συμβολισμούς, μπορούμε να γράψουμε στη γενική περίπτωση :

$$\mathbf{G}\left[\mathbf{m}_p + \mathbf{m}_0\right] = \left[\mathbf{d}_p + \mathbf{d}_0\right] \tag{6.14}$$

Το μήκος της λύσης είναι

$$L = \mathbf{m}^T\mathbf{m} = \left[\mathbf{m}_p + \mathbf{m}_0\right]^T \left[\mathbf{m}_p + \mathbf{m}_0\right] = \mathbf{m}_p^T\mathbf{m}_p + \mathbf{m}_0^T\mathbf{m}_0 \tag{6.15}$$

Το λάθος πρόβλεψης είναι

$$E = [\mathbf{d}_p + \mathbf{d}_0 - \mathbf{G}\mathbf{m}_p]^T[\mathbf{d}_p + \mathbf{d}_0 - \mathbf{G}\mathbf{m}_p] = \left[\mathbf{d}_p - \mathbf{G}\mathbf{m}_p\right]^T \left[\mathbf{d}_p - \mathbf{G}\mathbf{m}_p\right] + \mathbf{d}_0^T\mathbf{d}_0 \tag{6.16}$$

Όπου στους υπολογισμούς έχει ληφθεί υπ’ όψιν ότι γινόμενα διανυσμάτων που ανήκουν σε διαφορετικούς υπόχωρους είναι μηδενικά.

Σε ένα μεικτά ορισμένο πρόβλημα, εκ προοιμίου πληροφορία εισάγεται για να ορίσει παραμέτρους που περιέχονται στο χώρο $S_0(\mathbf{m})$ και το λάθος πρόβλεψης περιορίζεται στο χώρο $S_0(\mathbf{d})$ ικανοποιώντας τη σχέση $\mathbf{e}_p = \left[\mathbf{d}_p - \mathbf{G}\mathbf{m}_p\right] = 0$ επακριβώς.

Όταν το πρόβλημα είναι καθαρά υπο-ορισμένο, επιλέγοντας $m_0^{est} = 0$ (που την ονομάζουμε *φυσική λύση – natural solution*) οδηγούμαστε σε λύση ελαχίστου μήκους ενώ σε ένα καθαρά υπερορισμένο πρόβλημα οδηγούμαστε σε λύση ελαχίστων τετραγώνων.

6.3 Ανάλυση Ιδιαζουσών Τιμών (Singular Value Decomposition)

Στο κεφάλαιο αυτό θα αναφερθούμε σε μία διαδικασία που βοηθά στον καθορισμό του μηδενόχωρου και συνεπώς της λύσης ενός γραμμικού αντίστροφου προβλήματος όπως αυτή εκφράζεται από τη σχέση 6.3, μέσω κατάλληλης παραγοντοποίησης του πίνακα **G.**

6.3.1 Η ανάλυση

Θεωρούμε τον τετραγωνικό συμμετρικό πίνακα **S** διαστάσεων (*N*+*M*)Χ(*N*+*M*) που προκύπτει από τον πίνακα **G** του γραμμικού αντιστρόφου προβλήματος ως εξής :

$$\mathbf{S} = \begin{bmatrix} \mathbf{0} & \mathbf{G} \\ \mathbf{G^T} & \mathbf{0} \end{bmatrix} \tag{6.17}$$

Από τη Γραμμική Άλγεβρα γνωρίζομε ότι ο πίνακας αυτός έχει *N*+*M* πραγματικές ιδιοτιμές λ_i και μία πλήρη ομάδα ιδιοδιανυσμάτων $\mathbf{w}_i$ που επιλύουν το πρόβλημα

$$\mathbf{Sw_i} = \lambda_\mathbf{i}\mathbf{w_i}\,. \tag{6.18}$$

Διαχωρίζοντας το διάνυσμα $\mathbf{w}_i$ σε ένα μέρος διάστασης Ν και ένα άλλο διάστασης Μ, γράφομε την 6.18 αναλυτικά ως

$$\mathbf{Sw} = \begin{bmatrix} 0 & \mathbf{G} \\ \mathbf{G^T} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} = \lambda_i \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \tag{6.19}$$

Από την παραπάνω σχέση διαπιστώνουμε ότι ισχύουν ταυτόχρονα :

$$\mathbf{G}\mathbf{v}_i = \lambda_i \mathbf{u}_i \quad \text{και} \quad \mathbf{G}^T\mathbf{u}_i = \lambda_i \mathbf{v}_i\,. \tag{6.20}$$

Εάν υποθέσομε ότι έχομε μία θετική ιδιοτιμή λ_i τότε και η -λ_i είναι επίσης ιδιοτιμή με αντίστοιχα ιδιοδιανύσματα $[\mathbf{u}_i, \mathbf{v}_i]^T$ και $[-\mathbf{u}_i, \mathbf{v}_i]^T$. Εάν έχομε *p* θετικές ιδιοτιμές, θα πρέπει να έχομε συνολικά *N*+*M*-2*p* μηδενικές ιδιοτιμές. Φυσικά θα πρέπει να ισχύει ότι $p \le \min(N, M)$ αφού δεν μπορούμε να έχομε περισσότερες ιδιοτιμές σε σχέση με τη διάσταση του πίνακα.

Από τις 6.20 προκύπτουν εύκολα οι σχέσεις

$$\mathbf{G^T G}\mathbf{v}_i = \lambda_i^2 \mathbf{v}_i \quad \text{και} \quad \mathbf{G G^T}\mathbf{u}_i = \lambda_i^2 \mathbf{u}_i \tag{6.21}$$

Εξετάζοντας τους πίνακες $\mathbf{G^T G}$ και $\mathbf{G G^T}$ βλέπομε ότι είναι συμμετρικοί και διαθέτουν αντίστοιχα Μ και Ν ιδιοτιμές (με αντίστοιχα ιδιοδιανύσματα $\mathbf{v}_i$ και $\mathbf{u}_i$ που ορίζουν τα σύνολα **V** και **U** που έχουν τη μορφή πινάκων. Ο πίνακας **U** είναι :

$$\mathbf{U} = [\mathbf{u_1}, \mathbf{u_2}, \ldots\ldots \mathbf{u}_N] \tag{6.22}$$

και ο πίνακας **V** είναι

$$\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_M] \quad (6.23)$$

Σε κάθε σύνολο τα ιδιοδιανύσματα είναι ορθογώνια και συνεπώς παράγουν τους χώρους $S(\mathbf{m})$ και S(**d**) αντίστοιχα. Επίσης μπορούν να κανονικοποιηθούν ώστε να ισχύει :

$$\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}_N \text{ και } \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}_M \quad (6.24)$$

Με τους πίνακες $\mathbf{I}_N$ και $\mathbf{I}_M$ να είναι οι μοναδιαίοι πίνακες διαστάσεων N και M αντίστοιχα.

Ορίζοντας τώρα ως **Λ,** τον διαγώνιο πίνακα διαστάσεων NXM με διαγώνια στοιχεία τις μη αρνητικές ιδιοτιμές (τις χαρακτηρίζομε *ιδιάζουσες τιμές*) και παρατηρώντας ότι **V** και **U** είναι ορθογώνιοι, παίρνομε από την πρώτη από τις σχέσεις 6.20 :

$$\mathbf{GV}=\mathbf{U\Lambda} \quad (6.25)$$

και πολλαπλασιάζοντας από δεξιά με $\mathbf{V}^T$ παίρνομε :

$$\mathbf{G} = \mathbf{U\Lambda V}^T \quad (6.26)$$

που χαρακτηρίζετα ως παραγοντοποίηση ή *ανάλυση ιδιαζουσών τιμών* (*Singular Value Decomposition- SVD).*

6.3.2 Ανάλυση ιδιαζουσών τιμών και γενικευμένος αντίστροφος.

Όπως είπαμε πιο πάνω, οι ιδιάζουσες τιμές του πίνακα **G** δεν μπορεί να είναι περισσότερες από το $\min(N,M)$ συνεπώς στον πίνακα **Λ** μπορεί να έχομε γραμμές ή στήλες μηδενικές. Εάν N>M θα έχομε μία γραμμή τουλάχιστον με μηδενικά στοιχεία και εάν N<M θα έχομε τουλάχιστον μία στήλη με μηδενικά στοιχεία. Συνεπώς στη γενική περίπτωση θα μπορούσαμε να γράψομε τον πίνακα **Λ** στη μορφή

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (6.27)$$

Με τον πίνακα $\mathbf{\Lambda}_p$ να είναι ένας p X p διαγώνιος πίνακας με μη μηδενικά στοιχεία στη διαγώνιο.

Θα πρέπει να σημειώσουμε ότι η διάταξη των ιδιοτιμών στον πίνακα $\mathbf{\Lambda}_p$ γίνεται κατά φθίνουσα σειρά. Λαμβάνοντας υπ΄όψιν τη μορφή του **Λ**, η ανάλυση ιδιαζουσών τιμών (6.26) παίρνει τη μορφή :

$$\mathbf{G} = \mathbf{U}_p\mathbf{\Lambda}_p\mathbf{V}_p^{\,T} \quad (6.28)$$

με τους πίνακες $\mathbf{U}_p$ και $\mathbf{V}_p$ Να αποτελούνται από τις p πρώτες στήλες των $\mathbf{U}$ και $\mathbf{V}$ αντίστοιχα[2]. Ο πίνακας $\mathbf{G}$ δεν περιλαμβάνει συνεπώς πληροφορία για τους υπόχωρους που παράγονται από τις υπόλοιπες στήλες των $\mathbf{U}$ και $\mathbf{V}$ οι οποίες παράγουν τους πίνακες $\mathbf{U}_0$ και $\mathbf{V}_0$ αντίστοιχα.

Με άλλα λόγια καταφέραμε να παράγουμε το μηδενόχωρο $S_0(\mathbf{m})$ τα διανύσματα του οποίου είδαμε να συμμετέχουν στη γενική λύση του αντιστρόφου προβλήματος με την έννοια της (6.3) αλλά παράλληλα και τους υπόχωρους $S_p(\mathbf{m})$ και $S_p(\mathbf{d})$ που παράγονται από τα διανύσματα $\mathbf{V}_p$ και $\mathbf{U}_p$ αντίστοιχα.

Η *φυσική λύση* στο αντίστροφο πρόβλημα μπορεί τώρα να κατασκευαστεί μέσω της ανάλυσης ιδιαζουσών τιμών. Η λύση θα πρέπει να αποτελείται από ένα διάνυσμα $\mathbf{m}^{est}$ που δεν έχει στοιχεία στον $S_0(\mathbf{m})$ και ένα λάθος πρόβλεψης που δεν έχει στοιχεία στο $S_p(\mathbf{d})$.

Θεωρώντας τη λύση :

$$\mathbf{m}^{est} = \mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T \mathbf{d} \tag{6.29}$$

μπορούμε να αποδείξομε ότι ικανοποιεί τις παραπάνω απαιτήσεις και συνεπώς υιοθετείται ως η «**φυσική λύση**» του γραμμικού αντιστρόφου προβλήματος. Ο πίνακας $\mathbf{G}^{-g} = \mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T$ που προκύπτει από τη σχέση $\mathbf{m}^{est} = \mathbf{G}^{-g}\mathbf{d}$ χαρακτηρίζεται ως «ο φυσικός γενικευμένος αντίστροφος». Ο πίνακας ανάλυσης παραμέτρων για τον εν λόγω γενικευμένο αντίστροφο υπολογίζεται από τη σχέση της μορφής

$$\mathbf{R} = \mathbf{G}^{-g}\mathbf{G} = \{\mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T\}\{\mathbf{U}_p \mathbf{\Lambda}_p \mathbf{V}_p^T\} = \mathbf{V}_p \mathbf{V}_p^T \tag{6.30}$$

και ο πίνακας ανάλυσης δεδομένων από τη σχέση :

$$\mathbf{N} = \mathbf{G}\mathbf{G}^{-g} = \{\mathbf{U}_p \mathbf{\Lambda}_p \mathbf{V}_p^T\}\{\mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T\} = \mathbf{U}_p \mathbf{U}_p^T \tag{6.31}$$

Ο πίνακας συνδιακύμανσης για ασυσχέτιστα δεδομένα με ίδια διασπορά σ_d^2 είναι :

$$\left[\operatorname{cov}\mathbf{m}^{est}\right] = \mathbf{G}^{-g}\left[\operatorname{cov}\mathbf{d}\right]\mathbf{G}^{-gT} = \sigma_d^2\{\mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T\}\{\mathbf{U}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T\}^T = \sigma_d^2 \mathbf{V}_p \{\mathbf{\Lambda}_p^{-1}\}^2 \mathbf{V}_p^T \tag{6.32}$$

Είναι προφανές ότι τα δεδομένα (μετρήσεις) παράγονται πλήρως εάν ο υπόχωρος $\mathbf{U}_p$ παράγει το χώρο των μετρήσεων δηλαδή όταν $p=N$.

[2] Θυμηθείτε ότι υπάρχει διάταξη στις ιδιοτιμές συνεπώς πρέπει να υπάρξει και η αντιστοίχηση των ιδιοδιανυσμάτων.

Εάν υπάρχει εκ προοιμίου πληροφορία για τις παραμέτρους με μέση τιμή $\langle \mathbf{m} \rangle$ και πίνακα συνδιακύμανσης $[\mathrm{cov}\,\mathbf{m}]$ μπορεί να χρησιμοποιηθεί η σχέση

$$\mathbf{m}^{est} = \mathbf{G}^{-g}\mathbf{d} + [\mathbf{I} - \mathbf{R}]\langle \mathbf{m} \rangle \tag{6.33}$$

Ο πίνακας συνδιακύμανσης παίρνει τη μορφή

$$\left[\mathrm{cov}\,\mathbf{m}^{est}\right] = \mathbf{G}^{-g}[\mathrm{cov}\,\mathbf{d}]\mathbf{G}^{-gT} + [\mathbf{I} - \mathbf{R}][\mathrm{cov}\,\mathbf{m}][\mathbf{I} - \mathbf{R}]^{T} \tag{6.34}$$

Επισημαίνεται ότι η εφαρμογή της ανάλυσης ιδιαζουσών τιμών βασίζεται στη δυνατότητα ακριβούς υπολογισμού των τιμών αυτών. Σε πολλές περιπτώσεις όταν υπάρχουν μεγάλες διαφορές μεγέθους, που σημαίνει πρακτικά πίνακας **G** κακής κατάστασης, η ανάλυση δεν είναι ακριβής με τους υπάρχοντες αλγορίθμους. Σε άλλες περιπτώσεις ιδιάζουσες τιμές κοντά στο μηδέν θεωρούνται μηδενικές, οπότε στην πραγματικότητα επιλύεται ένα διαφορετικό πρόβλημα που προσεγγίζει το κανονικό. Άλλες λύσεις όπως για παράδειγμα η απόσβεση των μικρών ιδιοτιμών είναι επίσης εφαρμόσιμες.

Η γενική λύση που είδαμε με τη σχέση 6.3 μπορεί να θεωρηθεί ότι αποτελείται από το άθροισμα της φυσικής λύσης (μέσω της SVD) και της ειδικής που παράγεται από τα διανύσματα του μηδενόχωρου $S_0(\mathbf{m})$.