Krister Svanberg, april 2012

1 Positivt definita och positivt semidefinita matriser

Inom ickelinjär optimering, speciellt kvadratisk optimering, är det viktigt att på ett effektivt sätt kunna avgöra huruvida en given matris är positivt definit eller ej. Detta kapitel handlar om just detta, samt om en del relaterade ting. Vi inleder med lite fakta om några speciella typer av matriser.

1.1 Diagonalmatriser och triangulära matriser

En matris säges vara *kvadratisk* om den har lika många rader som kolonner.

En kvadratisk matris $\mathbf{H}$ säges vara *symmetrisk* om $\mathbf{H}^\mathsf{T} = \mathbf{H}$, dvs om elementen i $\mathbf{H}$ uppfyller $h_{ij} = h_{ji}$ för alla i och j.

En kvadratisk matris $\mathbf{D}$ säges vara *diagonal* om alla element som inte ligger i diagonalen är nollor, dvs om $d_{ij} = 0$ för alla i och j sådana att $i \neq j$. Diagonalelementen i en diagonalmatris $\mathbf{D}$ kallas ofta d_i (i stället för det mer korrekta d_{ii}). En diagonalmatris $\mathbf{D}$ är alltid symmetrisk, och den är icke-singulär om och endast om samtliga diagonalelement d_i är $\neq 0$.

En kvadratisk matris $\mathbf{L}$ säges vara *vänstertriangulär* om elementen i matrisen uppfyller att $l_{ij} = 0$ för alla i och j sådana att $i < j$, dvs om samtliga matriselement "nordost" om huvuddiagonalen är nollor. Den engelska termen är *lower triangular*. En vänstertriangulär matris $\mathbf{L}$ är icke-singulär om och endast om samtliga diagonalelement l_{ii} är $\neq 0$.

En kvadratisk matris $\mathbf{U}$ säges vara *högertriangulär* om elementen i matrisen uppfyller att $u_{ij} = 0$ för alla i och j sådana att $i > j$, dvs om samtliga matriselement "sydväst" om huvuddiagonalen är nollor. Den engelska termen är *upper triangular*. En högertriangulär matris $\mathbf{U}$ är icke-singulär om och endast om samtliga diagonalelement u_{ii} är $\neq 0$.

En 4×4 diagonalmatris $\mathbf{D}$, en 4×4 vänstertriangulär matris $\mathbf{L}$ och en 4×4 högertriangulär matris $\mathbf{U}$ ser alltså ut på följande vis:

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}$$

En viktig egenskap hos diagonala och triangulära matriser är att det är enkelt att lösa ekvationssystem där dessa är inblandade. Om man exempelvis ska lösa ett ekvationssystem $\mathbf{Lx} = \mathbf{b}$, med $\mathbf{L}$ enligt ovan, så får man lösningen $\mathbf{x}$ genom att i tur och ordning beräkna

$$x_1 = \frac{b_1}{l_{11}}, \quad x_2 = \frac{b_2 - l_{21}x_1}{l_{22}}, \quad x_3 = \frac{b_3 - l_{31}x_1 - l_{32}x_2}{l_{33}}, \quad x_4 = \frac{b_4 - l_{41}x_1 - l_{42}x_2 - l_{43}x_3}{l_{44}}.$$

Likaså om man ska lösa ett ekvationssystem $\mathbf{Ux} = \mathbf{b}$, med $\mathbf{U}$ enligt ovan, så får man lösningen $\mathbf{x}$ genom att i tur och ordning beräkna

$$x_4 = \frac{b_4}{u_{44}}, \quad x_3 = \frac{b_3 - u_{34}x_4}{u_{33}}, \quad x_2 = \frac{b_2 - u_{24}x_4 - u_{23}x_3}{u_{22}}, \quad x_1 = \frac{b_1 - u_{14}x_4 - u_{13}x_3 - u_{12}x_2}{u_{11}}.$$

Det allra lättaste är förstås att lösa ett ekvationssystem $\mathbf{Dx} = \mathbf{b}$, med $\mathbf{D}$ enligt ovan. Då får man lösningen $\mathbf{x}$ ur

$$x_1 = \frac{b_1}{d_1}\,, \quad x_2 = \frac{b_2}{d_2}\,, \quad x_3 = \frac{b_3}{d_3}\,, \quad x_4 = \frac{b_4}{d_4}\,.$$

1.2 Definition av positivt definit och positivt semidefinit matris

Låt $\mathbf{H} \in I\!R^{n\times n}$ vara en given symmetrisk matris och låt $\mathbf{x} \in I\!R^n$. Då är

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}x_i x_j. \tag{1.1}$$

Om $x_1, \ldots, x_n$ uppfattas som variabler medan matriselementen h_{ij} är konstanter,
så definierar (1.1) en *kvadratisk form* i $\mathbf{x} \in I\!R^n$.

En symmetrisk matris $\mathbf{H} \in I\!R^{n\times n}$ säges vara *positivt semidefinit* om

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} \geq 0 \;\text{ för alla } \mathbf{x} \in I\!R^n. \tag{1.2}$$

En symmetrisk matris $\mathbf{H} \in I\!R^{n\times n}$ säges vara *positivt definit* om

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} > 0 \;\text{ för alla } \mathbf{x} \in I\!R^n - \{\mathbf{0}\}. \tag{1.3}$$

Här betyder $\mathbf{x} \in I\!R^n - \{\mathbf{0}\}$ att vektorn $\mathbf{x} \in I\!R^n$ inte är nollvektorn. (Att man utesluter nollvektorn beror förstås på att om $\mathbf{x} = \mathbf{0}$ så är $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = 0$, oberoende av $\mathbf{H}$. Man kan alltså inte kräva att $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} > 0$ för *alla* $\mathbf{x} \in I\!R^n$, ty då skulle ingen matris $\mathbf{H}$ vara positivt definit!)
Observera att om $\mathbf{H}$ är positivt definit så är $\mathbf{H}$ per definition även positivt semidefinit.

Exempel 1. Antag att $\mathbf{H} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$. Med $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ är då

$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = x_1^2 + 2x_1x_2 + 2x_2x_1 + 5x_2^2 = x_1^2 + 4x_1x_2 + 5x_2^2 = (x_1 + 2x_2)^2 + x_2^2$,
som alltid är ≥ 0, med likhet om och endast om $x_1 + 2x_2 = 0$ och $x_2 = 0$,
dvs om och endast om $x_1 = x_2 = 0$.
Matrisen $\mathbf{H}$ är alltså positivt definit, och därmed även positivt semidefinit.

Exempel 2. Antag att $\mathbf{H} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Med $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ är då

$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = x_1^2 + 2x_1x_2 + 2x_2x_1 + 4x_2^2 = x_1^2 + 4x_1x_2 + 4x_2^2 = (x_1 + 2x_2)^2$,
som alltid är ≥ 0. Men $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = 0$ för (exempelvis) $\mathbf{x} = (-2,\, 1)^\mathsf{T} \neq \mathbf{0}$.
Matrisen $\mathbf{H}$ är alltså positivt semidefinit men inte positivt definit.

Exempel 3. Antag att $\mathbf{H} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. Med $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ är då

$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = x_1^2 + 2x_1x_2 + 2x_2x_1 + 3x_2^2 = x_1^2 + 4x_1x_2 + 3x_2^2 = (x_1 + 2x_2)^2 - x_2^2$,
som är < 0 för (exempelvis) $\mathbf{x} = (-2,\, 1)^\mathsf{T}$.
Matrisen $\mathbf{H}$ är alltså inte positivt semidefinit, och därmed inte heller positivt definit.

1.3 Några egenskaper hos positivt definita matriser

1.) Om $\mathbf{H}$ är symmetrisk och positivt definit så är $\mathbf{H}$ icke-singulär.
Bevis: Om $\mathbf{H}$ är singulär så finns ett $\mathbf{x} \neq \mathbf{0}$ sådant att $\mathbf{H}\mathbf{x} = \mathbf{0}$,
och då är även $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = 0$, vilket strider mot att $\mathbf{H}$ är positivt definit.

2.) Om $\mathbf{H}$ är symmetrisk och positivt definit så är varje diagonalelement $h_{ii} > 0$.
Bevis: Om exempelvis $h_{11} \leq 0$ så är, med $\mathbf{x} = (1, 0, \ldots, 0)^\mathsf{T}$, $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = h_{11} \leq 0$.

3.) En diagonalmatris $\mathbf{D}$ är positivt definit om och endast om alla d_i är > 0.
Bevis: $\mathbf{x}^\mathsf{T}\mathbf{D}\mathbf{x} = \sum_i d_i x_i^2$ som är > 0 för alla $\mathbf{x} \neq \mathbf{0}$ om och endast om alla $d_i > 0$.

4.) Om $\mathbf{H} \in I\!R^{n\times n}$ är symmetrisk och positivt definit, medan $\mathbf{B} \in I\!R^{n\times k}$ har linjärt oberoende kolonner, så är även $\mathbf{G} = \mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B} \in I\!R^{k\times k}$ symmetrisk och positivt definit.

Bevis: $\mathbf{G}^\mathsf{T} = (\mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{H}^\mathsf{T}\mathbf{B} = \mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B} = \mathbf{G}$ visar att $\mathbf{G}$ är symmetrisk.
För ett godtyckligt $\mathbf{x} \in I\!R^k$ är $\mathbf{x}^\mathsf{T}\mathbf{G}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B}\mathbf{x} = (\mathbf{B}\mathbf{x})^\mathsf{T}\mathbf{H}(\mathbf{B}\mathbf{x}) \geq 0$, med likhet endast om $\mathbf{B}\mathbf{x} = \mathbf{0}$, dvs endast om $\mathbf{x} = \mathbf{0}$, vilket visar att $\mathbf{G}$ är positivt definit.

5.) Antag att $\mathbf{H} \in I\!R^{n\times n}$ är en symmetrisk blockmatris på formen $\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{O}^\mathsf{T} \\ \mathbf{O} & \mathbf{H}_{22} \end{bmatrix}$,
där matriserna $\mathbf{H}_{11} \in I\!R^{n_1\times n_1}$ och $\mathbf{H}_{22} \in I\!R^{n_2\times n_2}$ båda är symmetriska,
medan $\mathbf{O} \in I\!R^{n_2\times n_1}$ och $\mathbf{O}^\mathsf{T} \in I\!R^{n_1\times n_2}$ betecknar nollmatriser.
Då är $\mathbf{H}$ positivt definit om och endast om $\mathbf{H}_{11}$ och $\mathbf{H}_{22}$ båda är positivt definita.

Bevis: Antag först att $\mathbf{H}_{11}$ och $\mathbf{H}_{22}$ båda är positivt definita och tag ett godtyckligt $\mathbf{x} \in I\!R^n$.
Då kan $\mathbf{x}$ skrivas $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \in I\!R^n$, där $\mathbf{x}_1 \in I\!R^{n_1}$ och $\mathbf{x}_2 \in I\!R^{n_2}$ (ty $n = n_1 + n_2$).
Om $\mathbf{x} \neq \mathbf{0}$ så är minst en av $\mathbf{x}_1$ och $\mathbf{x}_2 \neq \mathbf{0}$, varvid $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = \mathbf{x}_1^\mathsf{T}\mathbf{H}_{11}\mathbf{x}_1 + \mathbf{x}_2^\mathsf{T}\mathbf{H}_{22}\mathbf{x}_2 > 0$,
vilket visar att $\mathbf{H}$ är positivt definit.
Antag nu att $\mathbf{H}_{11}$ inte är positivt definit.
Då finns ett $\mathbf{x}_1 \neq \mathbf{0}$ sådant att $\mathbf{x}_1^\mathsf{T}\mathbf{H}_{11}\mathbf{x}_1 \leq 0$. Med $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{pmatrix} \in I\!R^n$ blir då
$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = \mathbf{x}_1^\mathsf{T}\mathbf{H}_{11}\mathbf{x}_1 \leq 0$, vilket visar att $\mathbf{H}$ inte är positivt definit.
Antag slutligen att $\mathbf{H}_{22}$ inte är positivt definit.
Då finns ett $\mathbf{x}_2 \neq \mathbf{0}$ sådant att $\mathbf{x}_2^\mathsf{T}\mathbf{H}_{22}\mathbf{x}_2 \leq 0$. Med $\mathbf{x} = \begin{pmatrix} \mathbf{0} \\ \mathbf{x}_2 \end{pmatrix} \in I\!R^n$ blir då
$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = \mathbf{x}_2^\mathsf{T}\mathbf{H}_{22}\mathbf{x}_2 \leq 0$, vilket visar att $\mathbf{H}$ inte är positivt definit.

6.) En symmetrisk matris $\mathbf{H}$ är positivt definit om och endast om samtliga dess egenvärden är > 0. (Bevisas ej här.)

1.4 Några egenskaper hos positivt semidefinita matriser

Vi använder här förkortningen PSD för positivt semidefinit.

1.) Om $\mathbf{H}$ är symmetrisk och PSD så är varje diagonalelement $h_{ii} \geq 0$.
Bevis: Om exempelvis $h_{11} < 0$ så är, med $\mathbf{x} = (1, 0, \ldots, 0)^\mathsf{T}$, $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = h_{11} < 0$.

2.) Om $\mathbf{H}$ är symmetrisk och PSD och ett diagonalelement uppfyller $h_{ii} = 0$, så är $h_{ij} = h_{ji} = 0$ för alla j. *Bevis:* Om exempelvis $h_{11} = 0$ och $h_{12} = h_{21} \neq 0$ så är, med $\mathbf{x} = (x_1, 1, 0, \ldots, 0)^\mathsf{T}$, $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = h_{22} + 2h_{12}x_1$, som är < 0 för lämpligt valt värde på x_1.

3.) En diagonalmatris $\mathbf{D}$ är PSD om och endast om alla diagonalelement d_i är ≥ 0.
Bevis: $\mathbf{x}^\mathsf{T}\mathbf{D}\mathbf{x} = \sum_i d_i x_i^2$ som är ≥ 0 för alla $\mathbf{x}$ om och endast om alla $d_i \geq 0$.

4.) Antag att $\mathbf{H} \in I\!R^{n\times n}$ är symmetrisk och PSD, och att $\mathbf{B} \in I\!R^{n\times k}$.
Då är $\mathbf{G} = \mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B} \in I\!R^{k\times k}$ symmetrisk och PSD.

Bevis: $\mathbf{G}^\mathsf{T} = (\mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{H}^\mathsf{T}\mathbf{B} = \mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B} = \mathbf{G}$ visar att $\mathbf{G}$ är symmetrisk.
För varje $\mathbf{x} \in I\!R^k$ är $\mathbf{x}^\mathsf{T}\mathbf{G}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{B}^\mathsf{T}\mathbf{H}\mathbf{B}\mathbf{x} = (\mathbf{B}\mathbf{x})^\mathsf{T}\mathbf{H}(\mathbf{B}\mathbf{x}) \geq 0$, vilket visar att $\mathbf{G}$ är PSD.

5.) Antag att $\mathbf{H} \in I\!R^{n\times n}$ är en symmetrisk blockmatris på formen $\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{O}^\mathsf{T} \\ \mathbf{O} & \mathbf{H}_{22} \end{bmatrix}$,
där matriserna $\mathbf{H}_{11} \in I\!R^{n_1\times n_1}$ och $\mathbf{H}_{22} \in I\!R^{n_2\times n_2}$ båda är symmetriska,
medan $\mathbf{O} \in I\!R^{n_2\times n_1}$ och $\mathbf{O}^\mathsf{T} \in I\!R^{n_1\times n_2}$ betecknar nollmatriser.
Då är $\mathbf{H}$ positivt semidefinit om och endast om $\mathbf{H}_{11}$ och $\mathbf{H}_{22}$ båda är positivt semidefinita.

Bevis: Antag att $\mathbf{H}_{11}$ och $\mathbf{H}_{22}$ båda är PSD och tag ett godtyckligt $\mathbf{x} \in I\!R^n$.
Då kan $\mathbf{x}$ skrivas $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \in I\!R^n$, där $\mathbf{x}_1 \in I\!R^{n_1}$ och $\mathbf{x}_2 \in I\!R^{n_2}$ (ty $n = n_1 + n_2$).
Därmed är $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = \mathbf{x}_1^\mathsf{T}\mathbf{H}_{11}\mathbf{x}_1 + \mathbf{x}_2^\mathsf{T}\mathbf{H}_{22}\mathbf{x}_2 \geq 0$, vilket visar att $\mathbf{H}$ är PSD.

Antag nu att $\mathbf{H}_{11}$ inte är PSD. Då finns ett $\mathbf{x}_1$ sådant att $\mathbf{x}_1^\mathsf{T}\mathbf{H}_{11}\mathbf{x}_1 < 0$.
Med $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{pmatrix} \in I\!R^n$ blir då $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = \mathbf{x}_1^\mathsf{T}\mathbf{H}_{11}\mathbf{x}_1 < 0$, vilket visar att $\mathbf{H}$ inte är PSD.

Antag slutligen att $\mathbf{H}_{22}$ inte är PSD. Då finns ett $\mathbf{x}_2$ sådant att $\mathbf{x}_2^\mathsf{T}\mathbf{H}_{22}\mathbf{x}_2 < 0$.
Med $\mathbf{x} = \begin{pmatrix} \mathbf{0} \\ \mathbf{x}_2 \end{pmatrix} \in I\!R^n$ blir då $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = \mathbf{x}_2^\mathsf{T}\mathbf{H}_{22}\mathbf{x}_2 < 0$, vilket visar att $\mathbf{H}$ inte är PSD.

6.) En symmetrisk matris $\mathbf{H}$ är positivt semidefinit om och endast om samtliga dess egenvärden är ≥ 0. (Bevisas ej här.)

1.5 Matriserna $\mathbf{A}^\mathsf{T}\mathbf{A}$ och $\mathbf{A}\mathbf{A}^\mathsf{T}$

Låt $\mathbf{A} \in I\!\!R^{m\times n}$ vara en given matris, och sätt $\mathbf{H} = \mathbf{A}^\mathsf{T}\mathbf{A} \in I\!\!R^{n\times n}$ och $\mathbf{G} = \mathbf{A}\mathbf{A}^\mathsf{T} \in I\!\!R^{m\times m}$.

Då är både $\mathbf{H}$ och $\mathbf{G}$ symmetriska, ty
$\mathbf{H}^\mathsf{T} = (\mathbf{A}^\mathsf{T}\mathbf{A})^\mathsf{T} = \mathbf{A}^\mathsf{T}(\mathbf{A}^\mathsf{T})^\mathsf{T} = \mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{H}$ och $\mathbf{G}^\mathsf{T} = (\mathbf{A}\mathbf{A}^\mathsf{T})^\mathsf{T} = (\mathbf{A}^\mathsf{T})^\mathsf{T}\mathbf{A}^\mathsf{T} = \mathbf{A}\mathbf{A}^\mathsf{T} = \mathbf{G}$.

Vidare är både $\mathbf{H}$ och $\mathbf{G}$ positivt semidefinita, ty om $\mathbf{x}$ är en godtycklig vektor i $I\!\!R^n$ och $\mathbf{y}$ är en godtycklig vektor i $I\!\!R^m$ så blir

$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} \;=\; \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} \;=\; (\mathbf{A}\mathbf{x})^\mathsf{T}(\mathbf{A}\mathbf{x}) \;=\; |\mathbf{A}\mathbf{x}|^2 \;\geq\; 0,$ och

$\mathbf{y}^\mathsf{T}\mathbf{G}\mathbf{y} \;=\; \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{y} \;=\; (\mathbf{A}^\mathsf{T}\mathbf{y})^\mathsf{T}(\mathbf{A}^\mathsf{T}\mathbf{y}) \;=\; |\mathbf{A}^\mathsf{T}\mathbf{y}|^2 \;\geq\; 0.$

Antag att $\mathbf{A}$ har *linjärt oberoende kolonner*. Då är $\mathbf{A}\mathbf{x} \neq \mathbf{0}$ för varje $\mathbf{x} \in I\!\!R^n - \{\mathbf{0}\}$.
Det betyder att för varje $\mathbf{x} \in I\!\!R^n - \{\mathbf{0}\}$ är

$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} \;=\; \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} \;=\; (\mathbf{A}\mathbf{x})^\mathsf{T}(\mathbf{A}\mathbf{x}) \;=\; |\mathbf{A}\mathbf{x}|^2 \;>\; 0,$

vilket visar att $\mathbf{A}^\mathsf{T}\mathbf{A}$ är positivt definit om kolonnerna i $\mathbf{A}$ är linjärt oberoende.

Antag att $\mathbf{A}$ har *linjärt oberoende rader*. Då är $\mathbf{A}^\mathsf{T}\mathbf{y} \neq \mathbf{0}$ för varje $\mathbf{y} \in I\!\!R^m - \{\mathbf{0}\}$.
Det betyder att för varje $\mathbf{y} \in I\!\!R^m - \{\mathbf{0}\}$ är

$\mathbf{y}^\mathsf{T}\mathbf{G}\mathbf{y} \;=\; \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{y} \;=\; (\mathbf{A}^\mathsf{T}\mathbf{y})^\mathsf{T}(\mathbf{A}^\mathsf{T}\mathbf{y}) \;=\; |\mathbf{A}^\mathsf{T}\mathbf{y}|^2 \;>\; 0,$

vilket visar att $\mathbf{A}\mathbf{A}^\mathsf{T}$ är positivt definit om raderna i $\mathbf{A}$ är linjärt oberoende.

Följande samband kommer ofta till användning:

$$\begin{aligned}
\mathcal{N}(\mathbf{A}^\mathsf{T}\mathbf{A}) &= \mathcal{N}(\mathbf{A})\,,\\
\mathcal{N}(\mathbf{A}\mathbf{A}^\mathsf{T}) &= \mathcal{N}(\mathbf{A}^\mathsf{T})\,,\\
\mathcal{R}(\mathbf{A}^\mathsf{T}\mathbf{A}) &= \mathcal{R}(\mathbf{A}^\mathsf{T})\,,\\
\mathcal{R}(\mathbf{A}\mathbf{A}^\mathsf{T}) &= \mathcal{R}(\mathbf{A})\,.
\end{aligned} \tag{1.4}$$

Dessa visas på följande sätt:

$\mathbf{x} \in \mathcal{N}(\mathbf{A}) \;\Rightarrow\; \mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{x} \in \mathcal{N}(\mathbf{A}^\mathsf{T}\mathbf{A}).$

$\mathbf{x} \in \mathcal{N}(\mathbf{A}^\mathsf{T}\mathbf{A}) \;\Rightarrow\; \mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = 0 \;\Rightarrow\; |\mathbf{A}\mathbf{x}|^2 = 0 \;\Rightarrow\; \mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{x} \in \mathcal{N}(\mathbf{A}).$

$\mathbf{y} \in \mathcal{N}(\mathbf{A}^\mathsf{T}) \;\Rightarrow\; \mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0} \;\Rightarrow\; \mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0} \;\Rightarrow\; \mathbf{y} \in \mathcal{N}(\mathbf{A}\mathbf{A}^\mathsf{T}).$

$\mathbf{y} \in \mathcal{N}(\mathbf{A}\mathbf{A}^\mathsf{T}) \Rightarrow \mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0} \Rightarrow \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{y} = 0 \Rightarrow |\mathbf{A}^\mathsf{T}\mathbf{y}|^2 = 0 \Rightarrow \mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0} \Rightarrow \mathbf{y} \in \mathcal{N}(\mathbf{A}^\mathsf{T}).$

$\mathbf{x} \in \mathcal{R}(\mathbf{A}^\mathsf{T}) \;\Leftrightarrow\; \mathbf{x} \in \mathcal{N}(\mathbf{A})^\perp \;\Leftrightarrow\; \mathbf{x} \in \mathcal{N}(\mathbf{A}^\mathsf{T}\mathbf{A})^\perp \;\Leftrightarrow\; \mathbf{x} \in \mathcal{R}(\mathbf{A}^\mathsf{T}\mathbf{A}).$

$\mathbf{y} \in \mathcal{R}(\mathbf{A}) \;\Leftrightarrow\; \mathbf{y} \in \mathcal{N}(\mathbf{A}^\mathsf{T})^\perp \;\Leftrightarrow\; \mathbf{y} \in \mathcal{N}(\mathbf{A}\mathbf{A}^\mathsf{T})^\perp \;\Leftrightarrow\; \mathbf{y} \in \mathcal{R}(\mathbf{A}\mathbf{A}^\mathsf{T}).$

1.6 LDL^T-faktorisering av positivt definita matriser

Om man vill avgöra huruvida en given symmetrisk matris är positivt definit eller ej så bör man använda så kallad LDL^T-faktorisering, som bygger på följande sats.

Sats: En symmetrisk matris $\mathbf{H} \in I\!R^{n\times n}$ är positivt definit *om och endast om* det finns en vänstertriangulär matris $\mathbf{L} \in I\!R^{n\times n}$ med alla $\ell_{ii} = 1$ och en diagonalmatris $\mathbf{D} \in I\!R^{n\times n}$ med alla $d_i > 0$ sådana att $\mathbf{H} = \mathbf{LDL}^\mathsf{T}$.

Vi ska i detta avsnitt ge ett konstruktivt bevis av denna sats genom att beskriva en algoritm för LDL^T-faktorisering av positivt definita matriser.

I hela detta avsnitt används förkortningen PD = symmetrisk och positivt definit.

Först konstaterar vi att om $\mathbf{L}$ och $\mathbf{D}$ är enligt ovan så följer av egenskaperna i avsnitt 1.3 att $\mathbf{LDL}^\mathsf{T}$ är PD, eftersom en diagonalmatris med alla diagonalelement > 0 är PD, och en triangulär matris med alla diagonalelement $\neq 0$ har såväl linjärt oberoende kolonner som linjärt oberoende rader.

I resten av detta avsnitt ska vi förutsätta att $\mathbf{H}$ är PD, och visa hur man då kan bestämma $\mathbf{L}$ och $\mathbf{D}$ med egenskaper enligt ovan, sådana att $\mathbf{H} = \mathbf{LDL}^\mathsf{T}$.

För att göra framställningen mindre teknisk (och förhoppningsvis lättare att följa) nöjer vi oss med att beskriva metodiken för specialfallet att $\mathbf{H}$ är en given symmetrisk 4×4–matris. Generaliseringen till en godtycklig symmetrisk $n\times n$–matris torde sedan vara uppenbar.

Vårt mål är alltså att genomföra faktoriseringen $\mathbf{H} = \mathbf{LDL}^\mathsf{T}$, där

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & h_{44} \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix} \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix}.$$

Av orsaker som snart kommer att framgå döper vi först om elementen i $\mathbf{H}$ till $h_{ij}^{(1)}$.
Vi har alltså att

$$\mathbf{H} = \begin{bmatrix} h_{11}^{(1)} & h_{12}^{(1)} & h_{13}^{(1)} & h_{14}^{(1)} \\ h_{21}^{(1)} & h_{22}^{(1)} & h_{23}^{(1)} & h_{24}^{(1)} \\ h_{31}^{(1)} & h_{32}^{(1)} & h_{33}^{(1)} & h_{34}^{(1)} \\ h_{41}^{(1)} & h_{42}^{(1)} & h_{43}^{(1)} & h_{44}^{(1)} \end{bmatrix} \quad \text{där} \quad h_{ij}^{(1)} = h_{ji}^{(1)} \quad \text{för alla } i \text{ och } j.$$

Att $\mathbf{H}$ är positivt definit medför speciellt att elementet $h_{11}^{(1)} > 0$. Därför kan man dra multipler av första raden i $\mathbf{H}$ från de övriga raderna, så att samtliga element nedanför diagonalelementet $h_{11}^{(1)}$ i första kolonnen blir nollor.
Detta motsvarar att matrisen $\mathbf{H}$ multipliceras från vänster med matrisen $\mathbf{E}_1$ nedan.
Första raden i $\mathbf{H}$ påverkas inte av dessa radoperationer, så den första raden i matrisen $\mathbf{E}_1\mathbf{H}$ är identisk med den första raden i $\mathbf{H}$. Övriga rader påverkas däremot, så dess element får

heta $h_{ij}^{(2)}$. Man har då att

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -l_{21} & 1 & 0 & 0 \\ -l_{31} & 0 & 1 & 0 \\ -l_{41} & 0 & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_1\mathbf{H} = \begin{bmatrix} h_{11}^{(1)} & h_{12}^{(1)} & h_{13}^{(1)} & h_{14}^{(1)} \\ 0 & h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} \\ 0 & h_{32}^{(2)} & h_{33}^{(2)} & h_{34}^{(2)} \\ 0 & h_{42}^{(2)} & h_{43}^{(2)} & h_{44}^{(2)} \end{bmatrix}.$$

där $l_{i1} = \dfrac{h_{i1}^{(1)}}{h_{11}^{(1)}}$ och $h_{ij}^{(2)} = h_{ij}^{(1)} - l_{i1}h_{1j}^{(1)} = h_{ij}^{(1)} - \dfrac{h_{i1}^{(1)}h_{1j}^{(1)}}{h_{11}^{(1)}}$ för $i = 2, 3, 4$ och $j = 2, 3, 4$.

Därefter kan man dra multipler av första kolonnen i $\mathbf{E}_1\mathbf{H}$ från de övriga kolonnerna, så att samtliga element till höger om diagonalelementet $h_{11}^{(1)}$ i första raden blir nollor.
Detta motsvarar att matrisen $\mathbf{E}_1\mathbf{H}$ multipliceras från höger med matrisen $\mathbf{E}_1^{\mathsf{T}}$. (Att det är samma matris $\mathbf{E}_1$ som dyker upp en gång till, fast nu transponerad, beror på att matrisen $\mathbf{H}$ är symmetrisk så att $h_{1j}^{(1)} = h_{j1}^{(1)}$ för $j = 2, 3, 4$. Därför är det samma multipler som används vid kolonnoperationerna som vid radoperationerna.) Elementen $h_{ij}^{(2)}$ påverkas inte av dessa operationer eftersom första kolonnen i $\mathbf{E}_1\mathbf{H}$ enbart har nollor nedanför $h_{11}^{(1)}$. Då är alltså

$$\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}} = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} \\ 0 & h_{32}^{(2)} & h_{33}^{(2)} & h_{34}^{(2)} \\ 0 & h_{42}^{(2)} & h_{43}^{(2)} & h_{44}^{(2)} \end{bmatrix}.$$

Att $\mathbf{H}$ är PD medför att $\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}$ är PD, vilket medför att matrisen $\begin{bmatrix} h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} \\ h_{32}^{(2)} & h_{33}^{(2)} & h_{34}^{(2)} \\ h_{42}^{(2)} & h_{43}^{(2)} & h_{44}^{(2)} \end{bmatrix}$ är PD, vilket medför att $h_{22}^{(2)} > 0$.

Då kan man dra multipler av andra raden i $\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}$ från de två sista raderna, så att samtliga element nedanför diagonalelementet $h_{22}^{(2)}$ i andra kolonnen blir nollor.
Detta motsvarar att matrisen $\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}$ multipliceras från vänster med matrisen $\mathbf{E}_2$ nedan.
De två första raderna i $\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}$ påverkas inte av dessa radoperationer, Övriga rader påverkas däremot, så dess element får heta $h_{ij}^{(3)}$. Man har då att

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -l_{32} & 1 & 0 \\ 0 & -l_{42} & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}} = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} \\ 0 & 0 & h_{33}^{(3)} & h_{34}^{(3)} \\ 0 & 0 & h_{43}^{(3)} & h_{44}^{(3)} \end{bmatrix}$$

där $l_{i2} = \dfrac{h_{i2}^{(2)}}{h_{22}^{(2)}}$ och $h_{ij}^{(3)} = h_{ij}^{(2)} - l_{i2}h_{2j}^{(2)} = h_{ij}^{(2)} - \dfrac{h_{i2}^{(2)}h_{2j}^{(2)}}{h_{22}^{(2)}}$ för $i = 3, 4$ och $j = 3, 4$.

Därefter kan man dra multipler av andra kolonnen i $\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top$ från de två sista kolonnerna, så att samtliga element till höger om diagonalelementet $h_{22}^{(2)}$ i andra raden blir nollor.
Detta motsvarar att matrisen $\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top$ multipliceras från höger med matrisen $\mathbf{E}_2^\top$.
Elementen $h_{ij}^{(3)}$ påverkas inte av dessa operationer. Då är alltså

$$\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & 0 & 0 \\ 0 & 0 & h_{33}^{(3)} & h_{34}^{(3)} \\ 0 & 0 & h_{43}^{(3)} & h_{44}^{(3)} \end{bmatrix}$$

Att $\mathbf{H}$ är PD medför att $\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top$ är PD, vilket medför att matrisen $\begin{bmatrix} h_{33}^{(3)} & h_{34}^{(3)} \\ h_{43}^{(3)} & h_{44}^{(3)} \end{bmatrix}$ är PD, vilket medför att $h_{33}^{(3)} > 0$.

Då kan man dra en multipel av tredje raden i $\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top$ från fjärde raden, så att elementet nedanför diagonalelementet $h_{33}^{(3)}$ blir en nolla. Detta motsvarar att matrisen $\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top$ multipliceras från vänster med matrisen $\mathbf{E}_3$ nedan.
De tre första raderna i påverkas inte av denna radoperation, Fjärde raden påverkas däremot, så dess element får heta $h_{ij}^{(4)}$. Man har då att

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -l_{43} & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & 0 & 0 \\ 0 & 0 & h_{33}^{(3)} & h_{34}^{(3)} \\ 0 & 0 & 0 & h_{44}^{(4)} \end{bmatrix}$$

där $l_{43} = \dfrac{h_{43}^{(3)}}{h_{33}^{(3)}}$ och $h_{44}^{(4)} = h_{44}^{(3)} - l_{43}h_{34}^{(3)} = h_{44}^{(3)} - \dfrac{h_{43}^{(3)}h_{34}^{(3)}}{h_{33}^{(3)}}$.

Slutligen kan man dra en multipel av tredje kolonnen i $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top$ från den fjärde kolonnen, så att elementet till höger om diagonalelementet $h_{33}^{(3)}$ blir en nolla.
Detta motsvarar att matrisen $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top$ multipliceras från höger med matrisen $\mathbf{E}_3^\top$.
Elementet $h_{44}^{(4)}$ påverkas inte av denna operation. Då är alltså

$$\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top\mathbf{E}_3^\top = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & 0 & 0 \\ 0 & 0 & h_{33}^{(3)} & 0 \\ 0 & 0 & 0 & h_{44}^{(4)} \end{bmatrix}. \tag{1.5}$$

Att $\mathbf{H}$ är PD medför att $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top\mathbf{E}_3^\top$ är PD, vilket i sin tur medför att $h_{44}^{(4)} > 0$.

Låt diagonalmatrisen i (1.5) heta $\mathbf{D}$, dvs

$$\mathbf{D} = \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^{\mathsf{T}}\mathbf{E}_2^{\mathsf{T}}\mathbf{E}_3^{\mathsf{T}}. \tag{1.6}$$

Eftersom varje radoperationsmatris $\mathbf{E}_k$ är icke-singulär, så är (1.6) ekvivalent med att

$$\mathbf{H} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}\mathbf{D}\,(\mathbf{E}_3^{\mathsf{T}})^{-1}(\mathbf{E}_2^{\mathsf{T}})^{-1}(\mathbf{E}_1^{\mathsf{T}})^{-1}. \tag{1.7}$$

Trevligt nog är det enkelt att beräkna såväl inversmatriserna $\mathbf{E}_k^{-1}$ som produkten $\mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}$. Man får nämligen att

$$\mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & 0 & 1 & 0 \\ l_{41} & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{E}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & l_{42} & 0 & 1 \end{bmatrix}, \quad \mathbf{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & l_{43} & 1 \end{bmatrix},$$

$$\text{samt} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & 0 & 1 & 0 \\ l_{41} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & l_{42} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & l_{43} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix}.$$

Produkten $\mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}$ är alltså en nedåt vänstertriangulär matris med ettor på diagonalen. Denna matris kallar vi för $\mathbf{L}$, dvs

$$\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}. \tag{1.8}$$

Då är $\mathbf{L}^{\mathsf{T}} = (\mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1})^{\mathsf{T}} = (\mathbf{E}_3^{-1})^{\mathsf{T}}(\mathbf{E}_2^{-1})^{\mathsf{T}}(\mathbf{E}_1^{-1})^{\mathsf{T}} = (\mathbf{E}_3^{\mathsf{T}})^{-1}(\mathbf{E}_2^{\mathsf{T}})^{-1}(\mathbf{E}_1^{\mathsf{T}})^{-1}$,
och därför övergår (1.7) till

$$\mathbf{H} = \mathbf{L}\mathbf{D}\mathbf{L}^{\mathsf{T}}. \tag{1.9}$$

$\mathrm{LDL}^{\mathsf{T}}$-faktoriseringen av $\mathbf{H}$ är därmed klar.

1.7 Ett exempel på LDL^T-faktorisering

Vi ska avgöra om följande matris $\mathbf{H}$ är positivt definit, eller positivt semidefinit, eller ingetdera, genom att LDL^T-faktorisera den.

$$\mathbf{H} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

Addera först $\frac{1}{2}$ gånger rad 1 till rad 2 och addera sedan $\frac{1}{2}$ gånger kolonn 1 till kolonn 2.
Det ger

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

Addera nu $\frac{2}{3}$ gånger rad 2 till rad 3 och addera sedan $\frac{2}{3}$ gånger kolonn 2 till kolonn 3.
Det ger

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{2}{3} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T}\mathbf{E}_2^\mathsf{T} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & \frac{4}{3} & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix},$$

Addera nu $\frac{3}{4}$ gånger rad 3 till rad 4 och addera sedan $\frac{3}{4}$ gånger kolonn 3 till kolonn 4.
Det ger

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{3}{4} & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T}\mathbf{E}_2^\mathsf{T}\mathbf{E}_3^\mathsf{T} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & \frac{4}{3} & 0 \\ 0 & 0 & 0 & \frac{5}{4} \end{bmatrix}.$$

Därmed är LDL^T-faktoriseringen klar, och man har att

$$\mathbf{H} = \mathbf{L}\mathbf{D}\mathbf{L}^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & -\frac{2}{3} & 1 & 0 \\ 0 & 0 & -\frac{3}{4} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & \frac{4}{3} & 0 \\ 0 & 0 & 0 & \frac{5}{4} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 1 & -\frac{3}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Eftersom alla diagonalelement i $\mathbf{D}$ är > 0 kan man dra slutsatsen att $\mathbf{H}$ är positivt definit.

1.8 $\mathbf{LDL}^\mathsf{T}$-faktorisering och kvadratkomplettering

Det finns ett nära samband mellan LDL^T-faktorisering och gammal hederlig kvadratkomplettering. Vi ska här illustrera detta samband på exemplet i föregående avsnitt.

Låt $\mathbf{x} \in I\!R^4$ och bilda den kvadratiska formen

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = 2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2 - 2x_3x_4 + 2x_4^2. \tag{1.10}$$

Med hjälp av LDL^T-faktoriseringen erhålls att

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{D}\mathbf{L}^\mathsf{T}\mathbf{x} = (\mathbf{L}^\mathsf{T}\mathbf{x})^\mathsf{T}\mathbf{D}(\mathbf{L}^\mathsf{T}\mathbf{x}), \quad \text{där}$$

$$\mathbf{L}^\mathsf{T}\mathbf{x} = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 1 & -\frac{3}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 - \frac{1}{2}x_2 \\ x_2 - \frac{2}{3}x_3 \\ x_3 - \frac{3}{4}x_4 \\ x_4 \end{pmatrix}.$$

Därmed är

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = (\mathbf{L}^\mathsf{T}\mathbf{x})^\mathsf{T}\mathbf{D}(\mathbf{L}^\mathsf{T}\mathbf{x}) = \begin{pmatrix} x_1 - \frac{1}{2}x_2 \\ x_2 - \frac{2}{3}x_3 \\ x_3 - \frac{3}{4}x_4 \\ x_4 \end{pmatrix}^\mathsf{T} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & \frac{4}{3} & 0 \\ 0 & 0 & 0 & \frac{5}{4} \end{bmatrix} \begin{pmatrix} x_1 - \frac{1}{2}x_2 \\ x_2 - \frac{2}{3}x_3 \\ x_3 - \frac{3}{4}x_4 \\ x_4 \end{pmatrix}, \quad \text{dvs}$$

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = (\mathbf{L}^\mathsf{T}\mathbf{x})^\mathsf{T}\mathbf{D}(\mathbf{L}^\mathsf{T}\mathbf{x}) = 2(x_1 - \tfrac{1}{2}x_2)^2 + \tfrac{3}{2}(x_2 - \tfrac{2}{3}x_3)^2 + \tfrac{4}{3}(x_3 - \tfrac{3}{4}x_4)^2 + \tfrac{5}{4}x_4^2. \tag{1.11}$$

Med hjälp av LDL^T-faktorisering har alltså den kvadratiska formen $\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x}$ i (1.10) överförts till en summa av kvadrater. Ett alternativt sätt att utföra detta är med *kvadratkomplettering*. Här följer en beskrivning av hur detta går till för ovanstående exempel.

Först elimineras blandade termer som innehåller faktorn x_1 med hjälp av omskrivningen $2x_1^2 - 2x_1x_2 = 2(x_1^2 - x_1x_2) = 2((x_1 - \frac{1}{2}x_2)^2 - \frac{1}{4}x_2^2)$. Detta ger att

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = 2(x_1 - \tfrac{1}{2}x_2)^2 + \tfrac{3}{2}x_2^2 - 2x_2x_3 + 2x_3^2 - 2x_3x_4 + 2x_4^2.$$

Sedan elimineras blandade termer som innehåller faktorn x_2 med hjälp av omskrivningen $\frac{3}{2}x_2^2 - 2x_2x_3 = \frac{3}{2}(x_2^2 - \frac{4}{3}x_2x_3) = \frac{3}{2}((x_2 - \frac{2}{3}x_3)^2 - \frac{4}{9}x_3^2)$. Detta ger att

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = 2(x_1 - \tfrac{1}{2}x_2)^2 + \tfrac{3}{2}(x_2 - \tfrac{2}{3}x_3)^2 + \tfrac{4}{3}x_3^2 - 2x_3x_4 + 2x_4^2.$$

Slutligen elimineras blandade termer som innehåller faktorn x_3 med hjälp av omskrivningen $\frac{4}{3}x_3^2 - 2x_3x_4 = \frac{4}{3}(x_3^2 - \frac{3}{2}x_3x_4) = \frac{4}{3}((x_3 - \frac{3}{4}x_4)^2 - \frac{9}{16}x_4^2)$. Detta ger att

$$\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} = 2(x_1 - \tfrac{1}{2}x_2)^2 + \tfrac{3}{2}(x_2 - \tfrac{2}{3}x_3)^2 + \tfrac{4}{3}(x_3 - \tfrac{3}{4}x_4)^2 + \tfrac{5}{4}x_4^2, \tag{1.12}$$

vilket är exakt detsamma som (1.11).

1.9 LDL^T-faktorisering av positivt semidefinita matriser

I avsnitt 1.6 beskrevs hur man med hjälp av LDL^T-faktorisering kan avgöra huruvida en given symmetrisk matris är positivt definit eller ej. En naturlig följdfråga är då om man på ett motsvarande sätt kan avgöra huruvuda en given symmetrisk matris är positivt *semidefinit* eller ej. Svaret är att det kan man göra med hjälp av en modifierad LDL^T-faktorisering, där man tillåter diagonalelement $h_{ii}^{(i)}$ att vara $= 0$.

Sats: En symmetrisk matris $\mathbf{H} \in I\!R^{n\times n}$ är positivt semidefinit *om och endast om* det finns en vänstertriangulär matris $\mathbf{L} \in I\!R^{n\times n}$ med alla $\ell_{ii} = 1$ och en diagonalmatris $\mathbf{D} \in I\!R^{n\times n}$ med alla $d_i \geq 0$ sådana att $\mathbf{H} = \mathbf{LDL}^\mathsf{T}$.

Vi ska i detta avsnitt ge ett konstruktivt bevis av denna sats genom att beskriva en algoritm för LDL^T-faktorisering av positivt semidefinita matriser.

I hela detta avsnitt används förkortningen PSD = symmetrisk och positivt semidefinit.

Först konstaterar vi att om $\mathbf{L}$ och $\mathbf{D}$ är enligt ovan så följer av egenskaperna i avsnitt 1.4 att $\mathbf{LDL}^\mathsf{T}$ är PSD.

Antag fortsättningsvis att $\mathbf{H}$ är PSD, och antag för enkelhets skull att $n = 4$.

Då kan man försöka använda metodiken i avsnitt 1.6 för att LDL^T-faktorisera matrisen $\mathbf{H}$.
Det som kan inträffa om $\mathbf{H}$ inte är positivt definit är att något diagonalelement $h_{ii}^{(i)}$ blir ≤ 0.
Men eftersom $\mathbf{H}$ är PSD så kan inte $h_{ii}^{(i)} < 0$, enligt egenskaperna i avsnitt 1.4.
Alltså är det "värsta" som kan hända att något eller några $h_{ii}^{(i)} = 0$.
Antag exempelvis att $h_{22}^{(2)} = 0$. Eftersom $\mathbf{H}$ förutsatts vara PSD så följer då att $h_{23}^{(2)} = h_{32}^{(2)} = h_{24}^{(2)} = h_{42}^{(2)} = 0$, enligt egenskaperna i avsnitt 1.4.

Man har då kommit till följande läge, jämför avsnitt 1.6.

$$\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T} = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} \\ 0 & h_{32}^{(2)} & h_{33}^{(2)} & h_{34}^{(2)} \\ 0 & h_{42}^{(2)} & h_{43}^{(2)} & h_{44}^{(2)} \end{bmatrix} = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & h_{33}^{(2)} & h_{34}^{(2)} \\ 0 & 0 & h_{43}^{(2)} & h_{44}^{(2)} \end{bmatrix}.$$

Då låter man helt enkelt

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{varvid} \quad \mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T}\mathbf{E}_2^\mathsf{T} = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & 0 & 0 \\ 0 & 0 & h_{33}^{(3)} & h_{34}^{(3)} \\ 0 & 0 & h_{43}^{(3)} & h_{44}^{(3)} \end{bmatrix},$$

med $h_{22}^{(2)} = 0$ och $h_{ij}^{(3)} = h_{ij}^{(2)}$ för $i = 3, 4$ och $j = 3, 4$.

Detta uttryck kan jämföras med motsvarande uttryck i avsnitt 1.6. Skillnaden är att där var $h_{22}^{(2)} > 0$, medan här är $h_{22}^{(2)} = 0$. Dessutom är matrisen $\mathbf{E}_2$ nu en enhetsmatris, vilket den typiskt inte var i avsnitt 1.6 (om inte ℓ_{32} och ℓ_{42} båda råkade bli $= 0$).

Sedan fortsätter processen som i avsnitt 1.6. När man är färdig har man kommit till läget

$$\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\top\mathbf{E}_2^\top\mathbf{E}_3^\top = \begin{bmatrix} h_{11}^{(1)} & 0 & 0 & 0 \\ 0 & h_{22}^{(2)} & 0 & 0 \\ 0 & 0 & h_{33}^{(3)} & 0 \\ 0 & 0 & 0 & h_{44}^{(4)} \end{bmatrix} = \mathbf{D}, \tag{1.13}$$

där samtliga $h_{ii}^{(i)} \geq 0$.

Ur detta erhålls att $\mathbf{H} = \mathbf{L}\mathbf{D}\mathbf{L}^\top$, med $\mathbf{D}$ enligt ovan och $\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}$ som i likhet med i avsnitt 1.6 är på formen

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix}. \tag{1.14}$$

1.10 Ett nytt exempel på LDL^T-faktorisering

Vi ska avgöra om följande matris $\mathbf{H}$ är positivt definit, eller positivt semidefinit, eller ingetdera, genom att LDL^T-faktorisera den.

$$\mathbf{H} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Addera först 1 gånger rad 1 till rad 2 och addera sedan 1 gånger kolonn 1 till kolonn 2. Det ger

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Addera nu 1 gånger rad 2 till rad 3 och addera sedan 1 gånger kolonn 2 till kolonn 3. Det ger

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T}\mathbf{E}_2^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix},$$

Addera nu 1 gånger rad 3 till rad 4 och addera sedan 1 gånger kolonn 3 till kolonn 4. Det ger

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{och} \quad \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{H}\mathbf{E}_1^\mathsf{T}\mathbf{E}_2^\mathsf{T}\mathbf{E}_3^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Därmed är LDL^T-faktoriseringen klar, och man har att

$$\mathbf{H} = \mathbf{L}\mathbf{D}\mathbf{L}^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Eftersom alla diagonalelement i $\mathbf{D}$ är ≥ 0 kan man dra slutsatsen att $\mathbf{H}$ är positivt semidefinit, men inte positivt definit eftersom ett diagonalelement i $\mathbf{D}$ är $= 0$.