Krister Svanberg, mars 2015

1 Grundläggande kalkyler med vektorer och matriser

Trots att läsaren säkert redan behärskar grundläggande vektor- och matriskalkyler, ges här i Kapitel 1 en repetition om just detta. Det som möjligen kan kännas nytt är tolkningarna av vad som egentligen pågår bakom de formella kalkylerna. I avsnitt 1.8 (kapitlets höjdpunkt) tolkas multiplikation av två matriser på fyra olika sätt. I Kapitel 2 presenteras grunderna för partitionerade matriser, och i Kapitel 3 behandlas baser, inverser och besläktade ting.

1.1 Vektorer i $\mathbb{R}^n$ och linjärkombinationer av sådana

Mängden av reella tal betecknas $\mathbb{R}$.
Mängden $\mathbb{R}^n$ består av alla kolonnvektorer med n st reella komponenter.
Om $\mathbf{x} \in \mathbb{R}^n$ och de n st komponenterna i vektorn $\mathbf{x}$ kallas $x_1, x_2, \ldots, x_n$ så är alltså

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Man kan addera vektorer i $\mathbb{R}^n$ och man kan multiplicera en vektor i $\mathbb{R}^n$ med en skalär (dvs ett reellt tal). Om $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^n$ och $\alpha \in \mathbb{R}$ så gäller per definition att

$$\mathbf{x}+\mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix} \quad \text{och} \quad \alpha\cdot\mathbf{x} = \alpha\cdot\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

Om $\mathbf{x} \in \mathbb{R}^n$ så är $\mathbf{x}^\mathsf{T}$ en radvektor med samma komponenter som kolonnvektorn $\mathbf{x}$, dvs

$$\mathbf{x}^\mathsf{T} = (x_1 \;\; x_2 \;\; \cdots \;\; x_n).$$

$\mathbf{x}^\mathsf{T}$ brukar utläsas *x-transponat.*

Låt nu $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vara k st givna vektorer i $\mathbb{R}^n$, dvs $\mathbf{v}_i \in \mathbb{R}^n$ för $i = 1, \ldots, k$.
Vektorn $\mathbf{w} \in \mathbb{R}^n$ säges vara en *linjärkombination* av dessa vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_k$ om det finns k st skalärer (dvs reella tal) $\alpha_1, \ldots, \alpha_k$ sådana att

$$\mathbf{w} = \alpha_1\cdot\mathbf{v}_1 + \cdots + \alpha_k\cdot\mathbf{v}_k = \sum_{i=1}^{k} \alpha_i\cdot\mathbf{v}_i.$$

Ibland skrivs skalärerna i stället till höger om (kolonn-)vektorerna, varvid gångertecknet $\cdot$ utelämnas, dvs

$$\mathbf{w} = \mathbf{v}_1\alpha_1 + \cdots + \mathbf{v}_k\alpha_k = \sum_{i=1}^{k} \mathbf{v}_i\alpha_i.$$

Detta ändrar inte betydelsen på något sätt, utan är bara ett annat skrivsätt.

(Att det ofta är naturligare att skriva skalären α_i till *höger* om kolonnvektorn $\mathbf{v}_i$ beror på att man då kan uppfatta kolonnvektorn som en $n\times 1$ matris och skalären som en 1×1 matris. Räknereglerna för matrismultiplikation $\mathbf{BC}$ (se nedan) förskriver att antalet kolonner i $\mathbf{B}$ ska överensstämma med antalet rader i $\mathbf{C}$ och det är uppfyllt om $\mathbf{B} = \mathbf{v}_i$ är $n \times 1$ och $\mathbf{C} = \alpha_i$ är 1×1, dvs om skalären α_i står till *höger* om kolonnvektorn $\mathbf{v}_i$. Det är däremot *inte* uppfyllt om $\mathbf{B} = \alpha_i$ är 1×1 och $\mathbf{C} = \mathbf{v}_i$ är $n\times 1$, dvs om skalären α_i står till *vänster* om kolonnvektorn $\mathbf{v}_i$, varför vi i detta fall använder gångertecknet $\cdot$ för att markera att det inte är fråga om matrismultiplikation. Gångertecknet $\cdot$ används nämligen normalt inte vid matrismultiplikationer.)

Att kolonnvektorn $\mathbf{w}$ är en linjärkombination av kolonnvektorerna $\mathbf{v}_1, \ldots, \mathbf{v}_k$ är ekvivalent med att *radvektorn* $\mathbf{w}^\mathsf{T}$ är en linjärkombination av *radvektorerna* $\mathbf{v}_1^\mathsf{T}, \ldots, \mathbf{v}_k^\mathsf{T}$, dvs

$$\mathbf{w}^\mathsf{T} = \alpha_1 \mathbf{v}_1^\mathsf{T} + \cdots + \alpha_k \mathbf{v}_k^\mathsf{T} = \sum_{i=1}^{k} \alpha_i \mathbf{v}_i^\mathsf{T}.$$

(Här är det naturligt att skriva skalären α_i till *vänster* om radvektorn $\mathbf{v}_i^\mathsf{T}$. Om skalären uppfattas som en 1×1 matris och radvektorn som en $1\times n$ matris så är räknereglerna för matrismultiplikation uppfyllda och vi behöver inte använda gångertecknet.)

1.2 Multiplikation av radvektor och kolonnvektor (= skalärprodukt)

Om $\mathbf{x} \in \mathbb{R}^n$ och $\mathbf{y} \in \mathbb{R}^n$ så kan man multiplicera radvektorn $\mathbf{x}^\mathsf{T}$ med kolonnvektorn $\mathbf{y}$. Resultatet blir per definition det reella talet

$$\mathbf{x}^\mathsf{T}\mathbf{y} \;=\; (x_1 \;\; x_2 \;\; \cdots \;\; x_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \;=\; x_1 y_1 + x_2 y_2 + \;\cdots\; + x_n y_n \;=\; \sum_{j=1}^{n} x_j y_j. \tag{1.1}$$

Eftersom resultatet blir en skalär brukar $\mathbf{x}^\mathsf{T}\mathbf{y}$ kallas för *skalärprodukten* av $\mathbf{x}$ och $\mathbf{y}$.
Man kan notera att ordningen på $\mathbf{x}$ och $\mathbf{y}$ inte spelar någon roll, dvs att $\mathbf{y}^\mathsf{T}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{y}$.
Vidare följer ur definitionen att om även $\mathbf{z} \in \mathbb{R}^n$ så är $\mathbf{x}^\mathsf{T}(\mathbf{y}+\mathbf{z}) = \mathbf{x}^\mathsf{T}\mathbf{y} + \mathbf{x}^\mathsf{T}\mathbf{z}$ och $(\mathbf{x}+\mathbf{y})^\mathsf{T}\mathbf{z} = \mathbf{x}^\mathsf{T}\mathbf{z} + \mathbf{y}^\mathsf{T}\mathbf{z}$.

Speciellt kan man skalärmultiplicera en vektor $\mathbf{x} \in \mathbb{R}^n$ med sig själv, varvid erhålls

$$\mathbf{x}^\mathsf{T}\mathbf{x} \;=\; x_1^2 + x_2^2 + \;\cdots\; + x_n^2 \;=\; \sum_{j=1}^{n} x_j^2 \;=\; |\mathbf{x}|^2,$$

där $|\mathbf{x}|$ betecknar (euklidiska) *längden* på vektorn $\mathbf{x}$.

Två vektorer $\mathbf{x} \in \mathbb{R}^n$ och $\mathbf{y} \in \mathbb{R}^n$ säges vara *ortogonala* om $\mathbf{x}^\mathsf{T}\mathbf{y} = 0$, dvs om skalärprodukten av $\mathbf{x}$ och $\mathbf{y}$ är noll. Ibland säger man *vinkelräta* i stället för ortogonala, eftersom den geometriska tolkningen i $\mathbb{R}^2$ och $\mathbb{R}^3$ är att vinkeln mellan $\mathbf{x}$ och $\mathbf{y}$ är 90 grader om $\mathbf{x}^\mathsf{T}\mathbf{y} = 0$.

Den enda vektor som är ortogonal mot sig själv är nollvektorn, ty

$$\mathbf{x}^\mathsf{T}\mathbf{x} = 0 \;\;\Leftrightarrow\;\; x_1^2 + x_2^2 + \;\cdots\; + x_n^2 = 0 \;\;\Leftrightarrow\;\; x_1 = x_2 = \;\cdots\; = x_n = 0 \;\;\Leftrightarrow\;\; \mathbf{x} = \mathbf{0}.$$

1.3 Matriser i $\mathbb{R}^{m\times n}$

Mängden $\mathbb{R}^{m\times n}$ består av alla $m\times n$ matriser med reella element.
Om matrisen $\mathbf{A}\in\mathbb{R}^{m\times n}$ så är alltså

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

där $a_{ij}\in\mathbb{R}$ betecknar matriselementet i rad nr i och kolonn nr j.

Man kan addera matriser av samma storlek och man kan multiplicera en matris med en skalär. Om $\mathbf{A}\in\mathbb{R}^{m\times n}$, $\mathbf{B}\in\mathbb{R}^{m\times n}$ och $\alpha\in\mathbb{R}$ så gäller per definition att

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}+\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}=\begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

$$\text{och}\qquad \alpha\cdot\mathbf{A}=\alpha\cdot\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}=\begin{bmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{bmatrix}.$$

1.4 Multiplikation av matris och kolonnvektor

Om $\mathbf{A}\in\mathbb{R}^{m\times n}$ och $\mathbf{x}\in\mathbb{R}^n$ så kan man multiplicera $\mathbf{A}$ och $\mathbf{x}$ (i den ordningen).
Resultatet blir per definition följande kolonnvektor $\mathbf{Ax}\in\mathbb{R}^m$:

$$\mathbf{Ax}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n \end{pmatrix}. \tag{1.2}$$

Om den i:te komponenten i kolonnvektorn $\mathbf{Ax}$ betecknas $(\mathbf{Ax})_i$ så gäller alltså att

$$(\mathbf{Ax})_i = a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=\sum_{j=1}^n a_{ij}x_j,\quad \text{för } i=1,\ldots,m. \tag{1.3}$$

Denna matris-kolonnvektormultiplikation kan tolkas på (åtminstone) två olika sätt.

Den **första tolkningen** av $\mathbf{Ax}$ utgår från kolonnerna i $\mathbf{A}$.
Låt $\hat{\mathbf{a}}_j$ beteckna den j:te kolonnen i matrisen $\mathbf{A}$, dvs

$$\mathbf{A}=\left[\,\hat{\mathbf{a}}_1\ \ \hat{\mathbf{a}}_2\ \cdots\ \hat{\mathbf{a}}_n\,\right],\quad \text{där }\ \hat{\mathbf{a}}_j=\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}\in\mathbb{R}^m.$$

Ur (1.2) erhålls att

$$\mathbf{Ax} = \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{pmatrix} + \cdots + \begin{pmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n,$$

och eftersom högerledet här är $\hat{\mathbf{a}}_1 x_1 + \cdots + \hat{\mathbf{a}}_n x_n$ så kan man tolka (1.2) på följande sätt:

$$\mathbf{Ax} = \left[\, \hat{\mathbf{a}}_1 \;\; \hat{\mathbf{a}}_2 \;\; \cdots \;\; \hat{\mathbf{a}}_n \,\right] \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \hat{\mathbf{a}}_1 x_1 + \hat{\mathbf{a}}_2 x_2 + \cdots + \hat{\mathbf{a}}_n x_n = \sum_{j=1}^{n} \hat{\mathbf{a}}_j x_j. \tag{1.4}$$

Kolonnvektorn $\mathbf{Ax}$ är alltså en linjärkombination av kolonnerna i matrisen $\mathbf{A}$.

Den **andra tolkningen** av $\mathbf{Ax}$ utgår från raderna i $\mathbf{A}$.
Låt $\bar{\mathbf{a}}_i^{\mathsf{T}}$ beteckna den i:te raden i matrisen $\mathbf{A}$, dvs

$$\mathbf{A} = \begin{bmatrix} \bar{\mathbf{a}}_1^{\mathsf{T}} \\ \bar{\mathbf{a}}_2^{\mathsf{T}} \\ \vdots \\ \bar{\mathbf{a}}_m^{\mathsf{T}} \end{bmatrix}, \quad \text{där} \;\; \bar{\mathbf{a}}_i^{\mathsf{T}} = (a_{i1} \;\; a_{i2} \;\; \cdots \;\; a_{in}), \quad \text{dvs} \;\; \bar{\mathbf{a}}_i \in I\!R^n.$$

Ur (1.3) erhålls då att

$$(\mathbf{Ax})_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = \bar{\mathbf{a}}_i^{\mathsf{T}} \mathbf{x},$$

vilket betyder att den i:te komponenten i vektorn $\mathbf{Ax}$ är skalärprodukten av den i:te raden i matrisen $\mathbf{A}$ och vektorn $\mathbf{x}$. Man kan alltså tolka (1.2) på följande sätt:

$$\mathbf{Ax} = \begin{bmatrix} \bar{\mathbf{a}}_1^{\mathsf{T}} \\ \bar{\mathbf{a}}_2^{\mathsf{T}} \\ \vdots \\ \bar{\mathbf{a}}_m^{\mathsf{T}} \end{bmatrix} \mathbf{x} = \begin{pmatrix} \bar{\mathbf{a}}_1^{\mathsf{T}} \mathbf{x} \\ \bar{\mathbf{a}}_2^{\mathsf{T}} \mathbf{x} \\ \vdots \\ \bar{\mathbf{a}}_m^{\mathsf{T}} \mathbf{x} \end{pmatrix}. \tag{1.5}$$

1.5 Multiplikation av radvektor och matris

Om $\mathbf{A} \in I\!R^{m \times n}$ och $\mathbf{u} \in I\!R^m$ så kan man multiplicera radvektorn $\mathbf{u}^{\mathsf{T}} = (u_1 \;\; u_2 \; \ldots \; u_m)$ och matrisen $\mathbf{A}$ (i den ordningen). Resultatet blir en radvektor $\mathbf{u}^{\mathsf{T}}\mathbf{A}$ med n st komponenter.
Om den j:te komponenten i denna radvektor betecknas $(\mathbf{u}^{\mathsf{T}}\mathbf{A})_j$ så gäller per definition att

$$(\mathbf{u}^{\mathsf{T}}\mathbf{A})_j = u_1 a_{1j} + u_2 a_{2j} + \cdots + u_m a_{mj} = \sum_{i=1}^{m} u_i a_{ij}, \quad \text{för } j = 1, \ldots, n. \tag{1.6}$$

Även denna radvektor-matrismultiplikation kan tolkas på (åtminstone) två olika sätt.

Den **första tolkningen** av $\mathbf{u}^{\mathsf{T}}\mathbf{A}$ utgår från kolonnerna i $\mathbf{A}$.
Låt $\hat{\mathbf{a}}_j$ beteckna den j:te kolonnen i $\mathbf{A}$. Ur (1.6) erhålls då att

$$(\mathbf{u}^{\mathsf{T}}\mathbf{A})_j \;=\; u_1 a_{1j} + u_2 a_{2j} + \;\cdots\; + u_m a_{mj} \;=\; \mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_j, \tag{1.7}$$

vilket betyder att den j:te komponenten i radvektorn $\mathbf{u}^{\mathsf{T}}\mathbf{A}$ är skalärprodukten av vektorn $\mathbf{u}$ och den j:te kolonnen i matrisen $\mathbf{A}$. Man kan alltså tolka (1.6) på följande sätt:

$$\mathbf{u}^{\mathsf{T}}\mathbf{A} \;=\; \mathbf{u}^{\mathsf{T}} \left[\, \hat{\mathbf{a}}_1 \;\; \hat{\mathbf{a}}_2 \;\cdots\; \hat{\mathbf{a}}_n \,\right] \;=\; (\mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_1 \;\; \mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_2 \;\cdots\; \mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_n). \tag{1.8}$$

Den **andra tolkningen** av $\mathbf{u}^{\mathsf{T}}\mathbf{A}$ utgår från raderna i $\mathbf{A}$.
Låt $\bar{\mathbf{a}}_i^{\mathsf{T}}$ beteckna den i:te raden i $\mathbf{A}$. Ur (1.6) erhålls då att

$$\mathbf{u}^{\mathsf{T}}\mathbf{A} \;=\; u_1\bar{\mathbf{a}}_1^{\mathsf{T}} + u_2\bar{\mathbf{a}}_2^{\mathsf{T}} + \;\cdots\; + u_m\bar{\mathbf{a}}_m^{\mathsf{T}},$$

så att (1.6) kan tolkas på följande sätt:

$$\mathbf{u}^{\mathsf{T}}\mathbf{A} \;=\; (u_1 \;\; u_2 \;\ldots\; u_m) \begin{bmatrix} \bar{\mathbf{a}}_1^{\mathsf{T}} \\ \bar{\mathbf{a}}_2^{\mathsf{T}} \\ \vdots \\ \bar{\mathbf{a}}_m^{\mathsf{T}} \end{bmatrix} \;=\; u_1\bar{\mathbf{a}}_1^{\mathsf{T}} + u_2\bar{\mathbf{a}}_2^{\mathsf{T}} + \;\cdots\; + u_m\bar{\mathbf{a}}_m^{\mathsf{T}} \;=\; \sum_{i=1}^{m} u_i\bar{\mathbf{a}}_i^{\mathsf{T}}. \tag{1.9}$$

Radvektorn $\mathbf{u}^{\mathsf{T}}\mathbf{A}$ är alltså en linjärkombination av raderna i matrisen $\mathbf{A}$.

1.6 Multiplikation av radvektor, matris och kolonnvektor

I detta avsnitt skall visas att om $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{u} \in \mathbb{R}^m$ och $\mathbf{x} \in \mathbb{R}^n$ så är $(\mathbf{u}^{\mathsf{T}}\mathbf{A})\mathbf{x} = \mathbf{u}^{\mathsf{T}}(\mathbf{A}\mathbf{x})$.
Genom att i tur och ordning använda (1.8), (1.1) och (1.4) så erhålls att

$$(\mathbf{u}^{\mathsf{T}}\mathbf{A})\mathbf{x} \;=\; (\mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_1 \;\; \mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_2 \;\cdots\; \mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \;=\; \sum_{j=1}^{n} \mathbf{u}^{\mathsf{T}}\hat{\mathbf{a}}_j x_j \;=\; \mathbf{u}^{\mathsf{T}}(\mathbf{A}\mathbf{x}).$$

Alternativt kan man i tur och ordning använda (1.5), (1.1) och (1.9):

$$\mathbf{u}^{\mathsf{T}}(\mathbf{A}\mathbf{x}) \;=\; (u_1 \;\; u_2 \;\cdots\; u_m) \begin{pmatrix} \bar{\mathbf{a}}_1^{\mathsf{T}}\mathbf{x} \\ \bar{\mathbf{a}}_2^{\mathsf{T}}\mathbf{x} \\ \vdots \\ \bar{\mathbf{a}}_m^{\mathsf{T}}\mathbf{x} \end{pmatrix} \;=\; \sum_{i=1}^{m} u_i\bar{\mathbf{a}}_i^{\mathsf{T}}\mathbf{x} \;=\; (\mathbf{u}^{\mathsf{T}}\mathbf{A})\mathbf{x}.$$

Man kan alltså utan risk för tvetydighet skriva $\mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{x}$.

Av ovanstående följer också att skalären $\mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{x}$ kan skrivas som dubbelsumman

$$\mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{x} \;=\; \sum_{i=1}^{m}\sum_{j=1}^{n} u_i a_{ij} x_j \;. \tag{1.10}$$

1.7 Multiplikation av kolonnvektor och radvektor

Om $\mathbf{b} \in \mathbb{R}^m$ och $\mathbf{c} \in \mathbb{R}^n$ så kan man multiplicera kolonnvektorn $\mathbf{b}$ och radvektorn $\mathbf{c}^\mathsf{T}$ (i den ordningen). Resultatet blir per definition följande matris $\mathbf{b}\mathbf{c}^\mathsf{T} \in \mathbb{R}^{m\times n}$:

$$\mathbf{b}\mathbf{c}^\mathsf{T} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} (c_1 \;\; c_2 \;\; \cdots \;\; c_n) = \begin{bmatrix} b_1c_1 & b_1c_2 & \cdots & b_1c_n \\ b_2c_1 & b_2c_2 & \cdots & b_2c_n \\ \vdots & \vdots & & \vdots \\ b_mc_1 & b_mc_2 & \cdots & b_mc_n \end{bmatrix}. \tag{1.11}$$

I denna matris $\mathbf{b}\mathbf{c}^\mathsf{T}$ är varje rad en multipel av vektorn $\mathbf{c}^\mathsf{T}$, medan varje kolonn är en multipel av vektorn $\mathbf{b}$, ty enligt (1.11) är

$$\mathbf{b}\mathbf{c}^\mathsf{T} = \begin{bmatrix} b_1\mathbf{c}^\mathsf{T} \\ b_2\mathbf{c}^\mathsf{T} \\ \vdots \\ b_m\mathbf{c}^\mathsf{T} \end{bmatrix} \quad \text{och} \quad \mathbf{b}\mathbf{c}^\mathsf{T} = \begin{bmatrix} \mathbf{b}c_1 & \mathbf{b}c_2 & \cdots & \mathbf{b}c_n \end{bmatrix}.$$

1.8 Multiplikation av två matriser

Om $\mathbf{B} \in \mathbb{R}^{m\times k}$ och $\mathbf{C} \in \mathbb{R}^{k\times n}$ så kan man multiplicera $\mathbf{B}$ och $\mathbf{C}$ (i den ordningen). Resultatet blir en $m\times n$-matris $\mathbf{BC} \in \mathbb{R}^{m\times n}$. Om $(\mathbf{BC})_{ij}$ betecknar elementet i den i:te raden och j:te kolonnen i matrisen $\mathbf{BC}$ så gäller per definition att

$$(\mathbf{BC})_{ij} = \sum_{\ell=1}^{k} b_{i\ell}c_{\ell j}, \quad \text{för } i = 1, \ldots, m \text{ och } j = 1, \ldots, n. \tag{1.12}$$

Denna matris-matrismultiplikation kan tolkas på (åtminstone) fyra olika sätt.

Den **första tolkningen** utgår från de enskilda elementen i matrisen $\mathbf{BC}$. Eftersom högerledet i (1.12) är produkten av den i:te raden i $\mathbf{B}$ och den j:te kolonnen i $\mathbf{C}$, dvs skalärprodukten av $\bar{\mathbf{b}}_i \in \mathbb{R}^\ell$ och $\hat{\mathbf{c}}_j \in \mathbb{R}^\ell$, så kan $\mathbf{BC}$ skrivas

$$\mathbf{BC} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T} \\ \bar{\mathbf{b}}_2^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{c}}_1 & \hat{\mathbf{c}}_2 & \cdots & \hat{\mathbf{c}}_n \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T}\hat{\mathbf{c}}_1 & \bar{\mathbf{b}}_1^\mathsf{T}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{b}}_1^\mathsf{T}\hat{\mathbf{c}}_n \\ \bar{\mathbf{b}}_2^\mathsf{T}\hat{\mathbf{c}}_1 & \bar{\mathbf{b}}_2^\mathsf{T}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{b}}_2^\mathsf{T}\hat{\mathbf{c}}_n \\ \vdots & \vdots & & \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T}\hat{\mathbf{c}}_1 & \bar{\mathbf{b}}_m^\mathsf{T}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{b}}_m^\mathsf{T}\hat{\mathbf{c}}_n \end{bmatrix}. \tag{1.13}$$

Den **andra tolkningen** utgår från kolonnerna i $\mathbf{BC}$. Genom att jämföra en godtycklig kolonn i högerledet i (1.13) med högerledet i (1.5) så inses att den j:te kolonnen i matrisen $\mathbf{BC}$ är produkten av matrisen $\mathbf{B}$ och den j:te kolonnen i $\mathbf{C}$, dvs

$$\mathbf{BC} = \mathbf{B}\begin{bmatrix} \hat{\mathbf{c}}_1 & \hat{\mathbf{c}}_2 & \cdots & \hat{\mathbf{c}}_n \end{bmatrix} = \begin{bmatrix} \mathbf{B}\hat{\mathbf{c}}_1 & \mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \mathbf{B}\hat{\mathbf{c}}_n \end{bmatrix}. \tag{1.14}$$

Den **tredje tolkningen** utgår från raderna i $\mathbf{BC}$. Genom att jämföra en godtycklig rad i högerledet i (1.13) med högerledet i (1.8) så inses att den i:te raden i matrisen $\mathbf{BC}$ är produkten av den i:te raden i $\mathbf{B}$ och matrisen $\mathbf{C}$, dvs

$$\mathbf{BC} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T} \\ \bar{\mathbf{b}}_2^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T} \end{bmatrix} \mathbf{C} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T}\mathbf{C} \\ \bar{\mathbf{b}}_2^\mathsf{T}\mathbf{C} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T}\mathbf{C} \end{bmatrix}. \tag{1.15}$$

Den **fjärde tolkningen** bygger på att produkten av den ℓ:te kolonnen i $\mathbf{B}$ och den ℓ:te raden i $\mathbf{C}$ enligt (1.11) ges av följande matris:

$$\hat{\mathbf{b}}_\ell \bar{\mathbf{c}}_\ell^\mathsf{T} = \begin{pmatrix} b_{1\ell} \\ b_{2\ell} \\ \vdots \\ b_{m\ell} \end{pmatrix} (c_{\ell 1} \;\; c_{\ell 2} \;\; \cdots \;\; c_{\ell n}) = \begin{bmatrix} b_{1\ell}c_{\ell 1} & b_{1\ell}c_{\ell 2} & \cdots & b_{1\ell}c_{\ell n} \\ b_{2\ell}c_{\ell 1} & b_{2\ell}c_{\ell 2} & \cdots & b_{2\ell}c_{\ell n} \\ \vdots & \vdots & & \vdots \\ b_{m\ell}c_{\ell 1} & b_{m\ell}c_{\ell 2} & \cdots & b_{m\ell}c_{\ell n} \end{bmatrix}.$$

Om $(\hat{\mathbf{b}}_\ell \bar{\mathbf{c}}_\ell^\mathsf{T})_{ij}$ betecknar elementet i den i:te raden och j:te kolonnen i denna matris $\hat{\mathbf{b}}_\ell \bar{\mathbf{c}}_\ell^\mathsf{T}$ så är alltså $(\hat{\mathbf{b}}_\ell \bar{\mathbf{c}}_\ell^\mathsf{T})_{ij} = b_{i\ell}c_{\ell j}$, vilket medför att (1.12) kan skrivas $(\mathbf{BC})_{ij} = \sum_{\ell=1}^{k} (\hat{\mathbf{b}}_\ell \bar{\mathbf{c}}_\ell^\mathsf{T})_{ij}$.

Eftersom detta gäller för varje rad i och kolonn j så är $\mathbf{BC} = \sum_{\ell=1}^{k} \hat{\mathbf{b}}_\ell \bar{\mathbf{c}}_\ell^\mathsf{T}$.

Den fjärde tolkningen av produkten $\mathbf{BC}$ kan därmed uttryckas på följande vis:

$$\mathbf{BC} = \begin{bmatrix} \hat{\mathbf{b}}_1 & \hat{\mathbf{b}}_2 & \cdots & \hat{\mathbf{b}}_k \end{bmatrix} \begin{bmatrix} \bar{\mathbf{c}}_1^\mathsf{T} \\ \bar{\mathbf{c}}_2^\mathsf{T} \\ \vdots \\ \bar{\mathbf{c}}_k^\mathsf{T} \end{bmatrix} = \hat{\mathbf{b}}_1 \bar{\mathbf{c}}_1^\mathsf{T} + \hat{\mathbf{b}}_2 \bar{\mathbf{c}}_2^\mathsf{T} + \cdots + \hat{\mathbf{b}}_k \bar{\mathbf{c}}_k^\mathsf{T}. \tag{1.16}$$

1.9 Multiplikation av tre matriser

Om $\mathbf{A} \in \mathbb{R}^{m\times p}$, $\mathbf{B} \in \mathbb{R}^{p\times q}$ och $\mathbf{C} \in \mathbb{R}^{q\times n}$ så kan man enligt ovan multiplicera dels $\mathbf{A}$ och $\mathbf{B}$, dels $\mathbf{B}$ och $\mathbf{C}$. Till resultat erhålls matriserna $\mathbf{AB} \in \mathbb{R}^{m\times q}$ respektive $\mathbf{BC} \in \mathbb{R}^{p\times n}$.
Men det betyder att man i sin tur kan multiplicera dels $\mathbf{AB}$ och $\mathbf{C}$, dels $\mathbf{A}$ och $\mathbf{BC}$.
Till resultat erhålls då matriserna $(\mathbf{AB})\mathbf{C} \in \mathbb{R}^{m\times n}$ respektive $\mathbf{A}(\mathbf{BC}) \in \mathbb{R}^{m\times n}$.

Här ska nu visas att matrismultiplikation är associativ, dvs att

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}). \tag{1.17}$$

Med hjälp av (1.15) och den högra likheten i (1.13) erhålls att

$$(\mathbf{AB})\mathbf{C} = \begin{bmatrix} \bar{\mathbf{a}}_1^\mathsf{T}\mathbf{B} \\ \bar{\mathbf{a}}_2^\mathsf{T}\mathbf{B} \\ \vdots \\ \bar{\mathbf{a}}_m^\mathsf{T}\mathbf{B} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{c}}_1 & \hat{\mathbf{c}}_2 & \cdots & \hat{\mathbf{c}}_n \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{a}}_1^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_1 & \bar{\mathbf{a}}_1^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{a}}_1^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_n \\ \bar{\mathbf{a}}_2^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_1 & \bar{\mathbf{a}}_2^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{a}}_2^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_n \\ \vdots & \vdots & & \vdots \\ \bar{\mathbf{a}}_m^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_1 & \bar{\mathbf{a}}_m^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{a}}_m^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_n \end{bmatrix}.$$

Med hjälp av (1.14) och den högra likheten i (1.13) erhålls på motsvarande sätt att

$$\mathbf{A}(\mathbf{BC}) = \begin{bmatrix} \bar{\mathbf{a}}_1^\mathsf{T} \\ \bar{\mathbf{a}}_2^\mathsf{T} \\ \vdots \\ \bar{\mathbf{a}}_m^\mathsf{T} \end{bmatrix} \begin{bmatrix} \mathbf{B}\hat{\mathbf{c}}_1 & \mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \mathbf{B}\hat{\mathbf{c}}_n \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{a}}_1^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_1 & \bar{\mathbf{a}}_1^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{a}}_1^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_n \\ \bar{\mathbf{a}}_2^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_1 & \bar{\mathbf{a}}_2^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{a}}_2^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_n \\ \vdots & \vdots & & \vdots \\ \bar{\mathbf{a}}_m^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_1 & \bar{\mathbf{a}}_m^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{a}}_m^\mathsf{T}\mathbf{B}\hat{\mathbf{c}}_n \end{bmatrix}.$$

Som synes är de bägge högerleden ovan identiska.
Man kan alltså utan risk för tvetydighet skriva $\mathbf{ABC}$.

1.10 Transponering av matriser

Att *transponera* matrisen $\mathbf{A}$ innebär att man låter rader och kolonner byta plats med varann. Om $\mathbf{A}$ är en $m\times n$-matris med elementet a_{ij} i rad nr i och kolonn nr j, så är därmed den transponerade matrisen $\mathbf{A}^\mathsf{T}$ en $n\times m$-matris med elementet a_{ij} i *kolonn* nr i och *rad* nr j.

Om

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}_1 & \hat{\mathbf{a}}_2 & \cdots & \hat{\mathbf{a}}_n \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{a}}_1^\mathsf{T} \\ \bar{\mathbf{a}}_2^\mathsf{T} \\ \vdots \\ \bar{\mathbf{a}}_m^\mathsf{T} \end{bmatrix},$$

så är alltså

$$\mathbf{A}^\mathsf{T} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}_1^\mathsf{T} \\ \hat{\mathbf{a}}_2^\mathsf{T} \\ \vdots \\ \hat{\mathbf{a}}_n^\mathsf{T} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{a}}_1 & \bar{\mathbf{a}}_2 & \cdots & \bar{\mathbf{a}}_m \end{bmatrix}.$$

Den j:te kolonnen $\hat{\mathbf{a}}_j$ i $\mathbf{A}$ har alltså transponerats och blivit den j:te raden $\hat{\mathbf{a}}_j^\mathsf{T}$ i $\mathbf{A}^\mathsf{T}$, medan den i:te raden $\bar{\mathbf{a}}_i^\mathsf{T}$ i $\mathbf{A}$ har transponerats och blivit den i:te kolonnen $\bar{\mathbf{a}}_i$ i $\mathbf{A}^\mathsf{T}$.

Man kan notera att om man transponerar en matris två gånger så får man tillbaka den ursprungliga matrisen, dvs $(\mathbf{A}^\mathsf{T})^\mathsf{T} = \mathbf{A}$.

Om $\mathbf{B} \in I\!R^{m\times k}$ och $\mathbf{C} \in I\!R^{k\times n}$ så kan man multiplicera $\mathbf{B}$ och $\mathbf{C}$, varvid $\mathbf{BC} \in I\!R^{m\times n}$. För de transponerade matriserna gäller då att $\mathbf{C}^\mathsf{T} \in I\!R^{n\times k}$ och $\mathbf{B}^\mathsf{T} \in I\!R^{k\times m}$, vilket betyder att man kan multiplicera $\mathbf{C}^\mathsf{T}$ och $\mathbf{B}^\mathsf{T}$, varvid $\mathbf{C}^\mathsf{T}\mathbf{B}^\mathsf{T} \in I\!R^{n\times m}$.

Eftersom

$$\mathbf{B} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T} \\ \bar{\mathbf{b}}_2^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T} \end{bmatrix} \quad \text{och} \quad \mathbf{C} = \begin{bmatrix} \hat{\mathbf{c}}_1 & \hat{\mathbf{c}}_2 & \cdots & \hat{\mathbf{c}}_n \end{bmatrix},$$

så är

$$\mathbf{B}^\mathsf{T} = \begin{bmatrix} \bar{\mathbf{b}}_1 & \bar{\mathbf{b}}_2 & \cdots & \bar{\mathbf{b}}_m \end{bmatrix} \quad \text{och} \quad \mathbf{C}^\mathsf{T} = \begin{bmatrix} \hat{\mathbf{c}}_1^\mathsf{T} \\ \hat{\mathbf{c}}_2^\mathsf{T} \\ \vdots \\ \hat{\mathbf{c}}_n^\mathsf{T} \end{bmatrix}.$$

Därmed är

$$\mathbf{C}^\mathsf{T}\mathbf{B}^\mathsf{T} = \begin{bmatrix} \hat{\mathbf{c}}_1^\mathsf{T} \\ \hat{\mathbf{c}}_2^\mathsf{T} \\ \vdots \\ \hat{\mathbf{c}}_n^\mathsf{T} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{b}}_1 & \bar{\mathbf{b}}_2 & \cdots & \bar{\mathbf{b}}_m \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{c}}_1^\mathsf{T}\bar{\mathbf{b}}_1 & \hat{\mathbf{c}}_1^\mathsf{T}\bar{\mathbf{b}}_2 & \cdots & \hat{\mathbf{c}}_1^\mathsf{T}\bar{\mathbf{b}}_m \\ \hat{\mathbf{c}}_2^\mathsf{T}\bar{\mathbf{b}}_1 & \hat{\mathbf{c}}_2^\mathsf{T}\bar{\mathbf{b}}_2 & \cdots & \hat{\mathbf{c}}_2^\mathsf{T}\bar{\mathbf{b}}_m \\ \vdots & \vdots & & \vdots \\ \hat{\mathbf{c}}_n^\mathsf{T}\bar{\mathbf{b}}_1 & \hat{\mathbf{c}}_n^\mathsf{T}\bar{\mathbf{b}}_2 & \cdots & \hat{\mathbf{c}}_n^\mathsf{T}\bar{\mathbf{b}}_m \end{bmatrix},$$

medan

$$\mathbf{BC} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T} \\ \bar{\mathbf{b}}_2^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{c}}_1 & \hat{\mathbf{c}}_2 & \cdots & \hat{\mathbf{c}}_n \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T}\hat{\mathbf{c}}_1 & \bar{\mathbf{b}}_1^\mathsf{T}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{b}}_1^\mathsf{T}\hat{\mathbf{c}}_n \\ \bar{\mathbf{b}}_2^\mathsf{T}\hat{\mathbf{c}}_1 & \bar{\mathbf{b}}_2^\mathsf{T}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{b}}_2^\mathsf{T}\hat{\mathbf{c}}_n \\ \vdots & \vdots & & \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T}\hat{\mathbf{c}}_1 & \bar{\mathbf{b}}_m^\mathsf{T}\hat{\mathbf{c}}_2 & \cdots & \bar{\mathbf{b}}_m^\mathsf{T}\hat{\mathbf{c}}_n \end{bmatrix}.$$

Från detta ser man att

$$(\mathbf{BC})^\mathsf{T} = \mathbf{C}^\mathsf{T}\mathbf{B}^\mathsf{T}, \tag{1.18}$$

som säger oss att transponatet av två matrisers produkt är lika med produkten av matrisernas transponat fast i omvänd ordning!

Detta kan generaliseras till flera matriser. Antag att $\mathbf{A} \in I\!R^{m\times p}$, $\mathbf{B} \in I\!R^{p\times q}$ och $\mathbf{C} \in I\!R^{q\times n}$. Då är produkten $\mathbf{ABC}$ enligt tidigare väldefinierad och lika med $\mathbf{A}(\mathbf{BC})$.
Upprepad användning av (1.18) ger då att

$$(\mathbf{ABC})^\mathsf{T} = (\mathbf{A}(\mathbf{BC}))^\mathsf{T} = (\mathbf{BC})^\mathsf{T}\mathbf{A}^\mathsf{T} = (\mathbf{C}^\mathsf{T}\mathbf{B}^\mathsf{T})\mathbf{A}^\mathsf{T} = \mathbf{C}^\mathsf{T}\mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}, \tag{1.19}$$

så även för tre matriser gäller att transponatet av matrisernas produkt är lika med produkten av matrisernas transponat fast i omvänd ordning. Resonemanget kan enkelt upprepas för produkten av fyra matriser, och sedan fem matriser, etc.

Antag speciellt att $\mathbf{A} \in I\!\!R^{m \times n}$, $\mathbf{u} \in I\!\!R^m$ och $\mathbf{x} \in I\!\!R^n$. Då gäller enligt (1.18) att

$$(\mathbf{A}\mathbf{x})^\mathsf{T} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T} \quad \text{och} \quad (\mathbf{u}^\mathsf{T}\mathbf{A})^\mathsf{T} = \mathbf{A}^\mathsf{T}\mathbf{u}, \tag{1.20}$$

ty $\mathbf{x} \in I\!\!R^n$ kan uppfatts som en $n \times 1$-matris och $\mathbf{u} \in I\!\!R^m$ som en $m \times 1$-matris.

Vidare gäller enligt (1.19) att

$$(\mathbf{u}^\mathsf{T}\mathbf{A}\mathbf{x})^\mathsf{T} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{u}. \tag{1.21}$$

Men $\mathbf{u}^\mathsf{T}\mathbf{A}\mathbf{x}$ är en 1×1-matris, dvs en skalär, som inte påverkas av transponering. Därför är $(\mathbf{u}^\mathsf{T}\mathbf{A}\mathbf{x})^\mathsf{T} = \mathbf{u}^\mathsf{T}\mathbf{A}\mathbf{x}$, vilket tillsammans med (1.21) ger att

$$\mathbf{u}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{u}, \tag{1.22}$$

som innebär att skalärprodukten av vektorerna $\mathbf{u}$ och $\mathbf{A}\mathbf{x}$ alltid är lika med skalärprodukten av vektorerna $\mathbf{x}$ och $\mathbf{A}^\mathsf{T}\mathbf{u}$.

2 Partitionerade matriser, även kallade blockmatriser

Vi inleder med ett exempel.

Givet följande fyra matriser $\mathbf{T} \in \mathbb{R}^{p\times q}$, $\mathbf{U} \in \mathbb{R}^{p\times s}$, $\mathbf{V} \in \mathbb{R}^{r\times q}$ och $\mathbf{W} \in \mathbb{R}^{r\times s}$:

$$\mathbf{T} = \begin{bmatrix} t_{11} & \cdots & t_{1q} \\ \vdots & & \vdots \\ t_{p1} & \cdots & t_{pq} \end{bmatrix}, \mathbf{U} = \begin{bmatrix} u_{11} & \cdots & u_{1s} \\ \vdots & & \vdots \\ u_{p1} & \cdots & u_{ps} \end{bmatrix}, \mathbf{V} = \begin{bmatrix} v_{11} & \cdots & v_{1q} \\ \vdots & & \vdots \\ v_{r1} & \cdots & v_{rq} \end{bmatrix}, \mathbf{W} = \begin{bmatrix} w_{11} & \cdots & w_{1s} \\ \vdots & & \vdots \\ w_{r1} & \cdots & w_{rs} \end{bmatrix}.$$

Observera att $\mathbf{T}$ har lika många rader som $\mathbf{U}$, att $\mathbf{V}$ har lika många rader som $\mathbf{W}$, att $\mathbf{T}$ har lika många kolonner som $\mathbf{V}$, och att $\mathbf{U}$ har lika många kolonner som $\mathbf{W}$.

Då kan man bilda följande matris $\mathbf{A}$ med $p + r$ st rader och $q + s$ st kolonner:

$$\mathbf{A} = \begin{bmatrix} t_{11} & \cdots & t_{1q} & u_{11} & \cdots & u_{1s} \\ \vdots & & \vdots & \vdots & & \vdots \\ t_{p1} & \cdots & t_{pq} & u_{p1} & \cdots & u_{ps} \\ v_{11} & \cdots & v_{1q} & w_{11} & \cdots & w_{1s} \\ \vdots & & \vdots & \vdots & & \vdots \\ v_{r1} & \cdots & v_{rq} & w_{r1} & \cdots & w_{rs} \end{bmatrix}.$$

Den naturliga beteckningen på denna matris är

$$\mathbf{A} = \begin{bmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{bmatrix}.$$

Här startade vi med fyra matriser och byggde med hjälp av dessa upp en större matris. Omvänt kan man förstås starta med en “stor” matris och dela upp den i mindre matriser:

Givet en matris $\mathbf{A}$ med m st rader och n st kolonner, låt m_1, m_2, n_1 och n_2 vara positiva heltal sådana att $m_1 + m_2 = m$ och $n_1 + n_2 = n$.
Inför sedan de fyra matriserna $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, $\mathbf{A}_{21}$ och $\mathbf{A}_{22}$ enligt följande:

Låt $\mathbf{A}_{11}$ bestå av elementen i de *första* m_1 raderna och de *första* n_1 kolonnerna i $\mathbf{A}$.
Låt $\mathbf{A}_{12}$ bestå av elementen i de *första* m_1 raderna och de *sista* n_2 kolonnerna i $\mathbf{A}$.
Låt $\mathbf{A}_{21}$ bestå av elementen i de *sista* m_2 raderna och de *första* n_1 kolonnerna i $\mathbf{A}$.
Låt $\mathbf{A}_{22}$ bestå av elementen i de *sista* m_2 raderna och de *sista* n_2 kolonnerna i $\mathbf{A}$.

Då kan $\mathbf{A}$ med beteckning i enlighet med ovan skrivas

$$\mathbf{A} \; = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}.$$

Man säger att $\mathbf{A}$ är en *partitionerad* (uppdelad) matris, eftersom $\mathbf{A}$ är uppdelad i delmatriser. Alternativt säger man att $\mathbf{A}$ är en *blockmatris* bestående av blocken $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, $\mathbf{A}_{21}$ och $\mathbf{A}_{22}$.

En partitionerad matris (blockmatris) behöver inte vara uppdelade i just *fyra* delar (block). Allmänt ser en partitionerad matris ut på följande sätt, där M och N är givna positiva heltal:

$$\mathbf{A} \;=\; \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1N} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2N} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{M1} & \mathbf{A}_{M2} & \cdots & \mathbf{A}_{MN} \end{bmatrix}.$$

Här måste alla delmatriser $\mathbf{A}_{ij}$ med gemensamt förstaindex i, exempelvis $\mathbf{A}_{21},\ \mathbf{A}_{22}, \cdots, \mathbf{A}_{2N}$, ha lika många rader. Vidare måste alla delmatriser $\mathbf{A}_{ij}$ med gemensamt andraindex j, exempelvis $\mathbf{A}_{12},\ \mathbf{A}_{22}, \cdots, \mathbf{A}_{M2}$, ha lika många kolonner.

Att partitionera matriser är ofta mycket användbart för att få överblick och insikt om vad som egentligen pågår vid matriskalkyler. Vi kommer att se många exempel på det längre fram. Här ska nu närmast härledas några grundläggande räkneregler för partitionerade matriser, speciellt matrismultiplikation.

2.1 Produkten av en $2{\times}1$ blockmatris och en $1{\times}2$ blockmatris

Antag att $\mathbf{B} \in I\!\!R^{m\times k}$ och $\mathbf{C} \in I\!\!R^{k\times n}$, så att produkten $\mathbf{BC}$ är definierad.
Antag vidare att $\mathbf{B}$ och $\mathbf{C}$ är partitionerade enligt

$$\mathbf{B} = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_p^\mathsf{T} \\ \bar{\mathbf{b}}_{p+1}^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \quad \text{och} \quad \mathbf{C} = \begin{bmatrix} \hat{\mathbf{c}}_1 & \cdots & \hat{\mathbf{c}}_q & \hat{\mathbf{c}}_{q+1} & \cdots & \hat{\mathbf{c}}_n \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix},$$

där

$$\mathbf{B}_1 = \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_p^\mathsf{T} \end{bmatrix},\ \ \mathbf{B}_2 = \begin{bmatrix} \bar{\mathbf{b}}_{p+1}^\mathsf{T} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T} \end{bmatrix},\ \ \mathbf{C}_1 = \begin{bmatrix} \hat{\mathbf{c}}_1 & \cdots & \hat{\mathbf{c}}_q \end{bmatrix} \ \text{ och } \ \mathbf{C}_2 = \begin{bmatrix} \hat{\mathbf{c}}_{q+1} & \cdots & \hat{\mathbf{c}}_n \end{bmatrix}.$$

Med hjälp av (1.15) erhålls då att

$$\mathbf{BC} \;=\; \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \mathbf{C} \;=\; \begin{bmatrix} \bar{\mathbf{b}}_1^\mathsf{T}\mathbf{C} \\ \vdots \\ \bar{\mathbf{b}}_p^\mathsf{T}\mathbf{C} \\ \bar{\mathbf{b}}_{p+1}^\mathsf{T}\mathbf{C} \\ \vdots \\ \bar{\mathbf{b}}_m^\mathsf{T}\mathbf{C} \end{bmatrix} \;=\; \begin{bmatrix} \mathbf{B}_1\mathbf{C} \\ \mathbf{B}_2\mathbf{C} \end{bmatrix}. \tag{2.1}$$

Med hjälp av (1.14) erhålls i stället att

$$\mathbf{BC} \;=\; \mathbf{B}\left[\,\mathbf{C}_1 \;\; \mathbf{C}_2\,\right] \;=\; \left[\,\mathbf{B}\hat{\mathbf{c}}_1 \;\cdots\; \mathbf{B}\hat{\mathbf{c}}_q \;\; \mathbf{B}\hat{\mathbf{c}}_{q+1} \;\cdots\; \mathbf{B}\hat{\mathbf{c}}_n\,\right] \;=\; \left[\,\mathbf{BC}_1 \;\; \mathbf{BC}_2\,\right], \tag{2.2}$$

samt att

$$\mathbf{B}_1\left[\,\mathbf{C}_1 \;\; \mathbf{C}_2\,\right] \;=\; \left[\,\mathbf{B}_1\hat{\mathbf{c}}_1 \;\cdots\; \mathbf{B}_1\hat{\mathbf{c}}_q \;\; \mathbf{B}_1\hat{\mathbf{c}}_{q+1} \;\cdots\; \mathbf{B}_1\hat{\mathbf{c}}_n\,\right] \;=\; \left[\,\mathbf{B}_1\mathbf{C}_1 \;\; \mathbf{B}_1\mathbf{C}_2\,\right], \tag{2.3}$$

$$\mathbf{B}_2\left[\,\mathbf{C}_1 \;\; \mathbf{C}_2\,\right] \;=\; \left[\,\mathbf{B}_2\hat{\mathbf{c}}_1 \;\cdots\; \mathbf{B}_2\hat{\mathbf{c}}_q \;\; \mathbf{B}_2\hat{\mathbf{c}}_{q+1} \;\cdots\; \mathbf{B}_2\hat{\mathbf{c}}_n\,\right] \;=\; \left[\,\mathbf{B}_2\mathbf{C}_1 \;\; \mathbf{B}_2\mathbf{C}_2\,\right]. \tag{2.4}$$

Genom att kombinera räknereglerna (2.1), (2.3) och (2.4) erhålls därmed att

$$\mathbf{BC} \;=\; \begin{bmatrix} \mathbf{B}_1\mathbf{C} \\ \\ \mathbf{B}_2\mathbf{C} \end{bmatrix} \;=\; \begin{bmatrix} \mathbf{B}_1\,[\,\mathbf{C}_1 \;\; \mathbf{C}_2\,] \\ \\ \mathbf{B}_2\,[\,\mathbf{C}_1 \;\; \mathbf{C}_2\,] \end{bmatrix} \;=\; \begin{bmatrix} \mathbf{B}_1\mathbf{C}_1 & \mathbf{B}_1\mathbf{C}_2 \\ \\ \mathbf{B}_2\mathbf{C}_1 & \mathbf{B}_2\mathbf{C}_2 \end{bmatrix}.$$

Alltså: Om antalet kolonner i såväl $\mathbf{B}_1$ som $\mathbf{B}_2$ överensstämmer med antalet rader i såväl $\mathbf{C}_1$ som $\mathbf{C}_2$ så är

$$\begin{bmatrix} \mathbf{B}_1 \\ \\ \mathbf{B}_2 \end{bmatrix}\left[\,\mathbf{C}_1 \;\; \mathbf{C}_2\,\right] \;=\; \begin{bmatrix} \mathbf{B}_1\mathbf{C}_1 & \mathbf{B}_1\mathbf{C}_2 \\ \\ \mathbf{B}_2\mathbf{C}_1 & \mathbf{B}_2\mathbf{C}_2 \end{bmatrix}. \tag{2.5}$$

2.2 Produkten av en 1×2 blockmatris och en 2×1 blockmatris

Antag nu att $\mathbf{G} \in I\!R^{m\times k}$ och $\mathbf{H} \in I\!R^{k\times n}$, så att produkten $\mathbf{GH}$ är definierad.
Antag vidare att $\mathbf{G}$ och $\mathbf{H}$ är partitionerade enligt

$$\mathbf{G} = \left[\,\hat{\mathbf{g}}_1 \;\cdots\; \hat{\mathbf{g}}_\ell \;\; \hat{\mathbf{g}}_{\ell+1} \;\cdots\; \hat{\mathbf{g}}_k\,\right] = \left[\,\mathbf{G}_1 \;\; \mathbf{G}_2\,\right] \quad \text{och} \quad \mathbf{H} = \begin{bmatrix} \bar{\mathbf{h}}_1^\mathsf{T} \\ \vdots \\ \bar{\mathbf{h}}_\ell^\mathsf{T} \\ \bar{\mathbf{h}}_{\ell+1}^\mathsf{T} \\ \vdots \\ \bar{\mathbf{h}}_k^\mathsf{T} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \\ \mathbf{H}_2 \end{bmatrix}, \;\; \text{där}$$

$$\mathbf{G}_1 = \left[\,\hat{\mathbf{g}}_1 \;\cdots\; \hat{\mathbf{g}}_\ell\,\right], \;\; \mathbf{G}_2 = \left[\,\hat{\mathbf{g}}_{\ell+1} \;\cdots\; \hat{\mathbf{g}}_k\,\right], \;\; \mathbf{H}_1 = \begin{bmatrix} \bar{\mathbf{h}}_1^\mathsf{T} \\ \vdots \\ \bar{\mathbf{h}}_\ell^\mathsf{T} \end{bmatrix} \;\; \text{och} \;\; \mathbf{H}_2 = \begin{bmatrix} \bar{\mathbf{h}}_{\ell+1}^\mathsf{T} \\ \vdots \\ \bar{\mathbf{h}}_k^\mathsf{T} \end{bmatrix}.$$

Med hjälp av (1.16) erhålls då att

$$\mathbf{GH} \;=\; \hat{\mathbf{g}}_1\bar{\mathbf{h}}_1^\mathsf{T} + \cdots + \hat{\mathbf{g}}_\ell\bar{\mathbf{h}}_\ell^\mathsf{T} \;+\; \hat{\mathbf{g}}_{\ell+1}\bar{\mathbf{h}}_{\ell+1}^\mathsf{T} + \cdots + \hat{\mathbf{g}}_k\bar{\mathbf{h}}_k^\mathsf{T} \;=\; \mathbf{G}_1\mathbf{H}_1 + \mathbf{G}_2\mathbf{H}_2\,.$$

Alltså: Om antalet rader i $\mathbf{G}_1$ = antalet rader i $\mathbf{G}_2$,
antalet kolonner i $\mathbf{H}_1$ = antalet kolonner i $\mathbf{H}_2$,
antalet kolonner i $\mathbf{G}_1$ = antalet rader i $\mathbf{H}_1$ och
antalet kolonner i $\mathbf{G}_2$ = antalet rader i $\mathbf{H}_2$, så är

$$\begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} = \mathbf{G}_1\mathbf{H}_1 + \mathbf{G}_2\mathbf{H}_2 \,. \tag{2.6}$$

2.3 Produkten av en 2×2 blockmatris och en 2×2 blockmatris

Antag nu att

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \quad \text{och} \quad \mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix},$$

där $\mathbf{B}_{ij} \in \mathbb{R}^{m_i \times k_j}$ och $\mathbf{C}_{ij} \in \mathbb{R}^{k_i \times n_j}$, för $i \in \{1,2\}$ och $j \in \{1,2\}$, så att
$\mathbf{B} \in \mathbb{R}^{m\times k}$ och $\mathbf{C} \in \mathbb{R}^{k\times n}$, där $m = m_1+m_2$, $n = n_1+n_2$ och $k = k_1+k_2$.
Nu ska härledas en räkneregel för produkten $\mathbf{BC}$.

Låt $\mathbf{G}_1 = \begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix}$, $\mathbf{G}_2 = \begin{bmatrix} \mathbf{B}_{12} \\ \mathbf{B}_{22} \end{bmatrix}$, $\mathbf{H}_1 = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \end{bmatrix}$ och $\mathbf{H}_2 = \begin{bmatrix} \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}$.

Då är $\mathbf{B} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 \end{bmatrix}$ och $\mathbf{C} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix}$, varvid räkneregeln (2.6) ger att

$$\mathbf{BC} = \mathbf{G}_1\mathbf{H}_1 + \mathbf{G}_2\mathbf{H}_2 = \begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_{12} \\ \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}. \tag{2.7}$$

Tillämpning av räkneregeln (2.5) på högerledet i (2.7) ger att

$$\mathbf{BC} = \begin{bmatrix} \mathbf{B}_{11}\mathbf{C}_{11} & \mathbf{B}_{11}\mathbf{C}_{12} \\ \mathbf{B}_{21}\mathbf{C}_{11} & \mathbf{B}_{21}\mathbf{C}_{12} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_{12}\mathbf{C}_{21} & \mathbf{B}_{12}\mathbf{C}_{22} \\ \mathbf{B}_{22}\mathbf{C}_{21} & \mathbf{B}_{22}\mathbf{C}_{22} \end{bmatrix}.$$

Alltså: Förutsatt att blocken $\mathbf{B}_{ij}$ och $\mathbf{C}_{ij}$ har dimensioner enligt ovan så är

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{11}\mathbf{C}_{11} + \mathbf{B}_{12}\mathbf{C}_{21} & \mathbf{B}_{11}\mathbf{C}_{12} + \mathbf{B}_{12}\mathbf{C}_{22} \\ \mathbf{B}_{21}\mathbf{C}_{11} + \mathbf{B}_{22}\mathbf{C}_{21} & \mathbf{B}_{21}\mathbf{C}_{12} + \mathbf{B}_{22}\mathbf{C}_{22} \end{bmatrix}. \tag{2.8}$$

Om man jämför (2.8) med multiplikationsregeln för "vanliga" 2×2 matriser så ser man att blockmatriser kan multipliceras med varann genom att helt enkelt behandla blocken som element vid vanlig matrismultiplikation. Detta förutsätter dock att dimensionerna på blocken är sådana att de inblandade blockmultiplikationerna och blockadditionerna är definierade. Denna viktiga räkneregel gäller även för mer generella blockmatriser än 2×2 blockmatriser.

2.4 Transponering av blockmatriser

Antag att $\mathbf{A}$ är följande 2×2 blockmatris:

$$\mathbf{A} = \begin{bmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{bmatrix} = \begin{bmatrix} t_{11} & \cdots & t_{1q} & u_{11} & \cdots & u_{1s} \\ \vdots & & \vdots & \vdots & & \vdots \\ t_{p1} & \cdots & t_{pq} & u_{p1} & \cdots & u_{ps} \\ v_{11} & \cdots & v_{1q} & w_{11} & \cdots & w_{1s} \\ \vdots & & \vdots & \vdots & & \vdots \\ v_{r1} & \cdots & v_{rq} & w_{r1} & \cdots & w_{rs} \end{bmatrix}.$$

Om man här transponerar $\mathbf{A}$, dvs byter plats på rader och kolonner, så erhålls

$$\mathbf{A}^\mathsf{T} = \begin{bmatrix} t_{11} & \cdots & t_{p1} & v_{11} & \cdots & v_{r1} \\ \vdots & & \vdots & \vdots & & \vdots \\ t_{1q} & \cdots & t_{pq} & v_{1q} & \cdots & v_{rq} \\ u_{11} & \cdots & u_{p1} & w_{11} & \cdots & w_{r1} \\ \vdots & & \vdots & \vdots & & \vdots \\ u_{1s} & \cdots & u_{ps} & w_{1s} & \cdots & w_{rs} \end{bmatrix},$$

vilket betyder att

$$\mathbf{A}^\mathsf{T} = \begin{bmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{bmatrix}^\mathsf{T} = \begin{bmatrix} \mathbf{T}^\mathsf{T} & \mathbf{V}^\mathsf{T} \\ \mathbf{U}^\mathsf{T} & \mathbf{W}^\mathsf{T} \end{bmatrix}.$$

Observera att blocken $\mathbf{U}$ och $\mathbf{V}$ byter plats och att samtliga block transponeras.

Allmänt gäller att om

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1N} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2N} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{M1} & \mathbf{A}_{M2} & \cdots & \mathbf{A}_{MN} \end{bmatrix}.$$

så är

$$\mathbf{A}^\mathsf{T} = \begin{bmatrix} \mathbf{A}_{11}^\mathsf{T} & \mathbf{A}_{21}^\mathsf{T} & \cdots & \mathbf{A}_{M1}^\mathsf{T} \\ \mathbf{A}_{12}^\mathsf{T} & \mathbf{A}_{22}^\mathsf{T} & \cdots & \mathbf{A}_{M2}^\mathsf{T} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{1N}^\mathsf{T} & \mathbf{A}_{2N}^\mathsf{T} & \cdots & \mathbf{A}_{MN}^\mathsf{T} \end{bmatrix}.$$

3 Baser och inverser

3.1 Linjärt beroende resp linjärt oberoende vektorer i $I\!R^m$

Låt $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vara n st givna vektorer i $I\!R^m$ (dvs $\mathbf{a}_i \in I\!R^m$ för $i = 1, \ldots, n$). Dessa vektorer säges vara *linjärt beroende* om det finns n st (reella) tal $x_1, \ldots, x_n$ som inte alla är 0, sådana att

$$\mathbf{a}_1 x_1 + \cdots + \mathbf{a}_n x_n = \mathbf{0}, \tag{3.1}$$

där $\mathbf{0} = (0, \ldots, 0)^\mathsf{T} \in I\!R^m$.

Om däremot den enda möjligheten att uppfylla (3.1) är att sätta $x_1 = \cdots = x_n = 0$ så säges vektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vara *linjärt oberoende.*

Låt $\mathbf{A}$ vara en $m \times n$ matris med de givna vektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_n$ till kolonner, dvs $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$, och betrakta följande homogena ekvationssystem i variabelvektorn $\mathbf{x} = (x_1, \ldots, x_n)^\mathsf{T} \in I\!R^n$:

$$\mathbf{A}\mathbf{x} = \mathbf{0}. \tag{3.2}$$

Eftersom (3.2) bara är ett alternativt sätt att skriva (3.1) på, så kan ovanstående definitioner ekvivalent uttryckas så här:
Kolonnerna till matrisen $\mathbf{A}$ är linjärt oberoende om det homogena systemet (3.2) endast har lösningen $\mathbf{x} = \mathbf{0}$.
Om det däremot finns minst en lösning $\mathbf{x} \neq \mathbf{0}$ till det homogena systemet (3.2) så är kolonnerna till $\mathbf{A}$ linjärt beroende.

Betrakta nu följande ekvationssystem, där $\mathbf{b} \in I\!R^m$ är en given högerledsvektor.

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{3.3}$$

Om kolonnerna i $\mathbf{A}$ är linjärt oberoende så har detta ekvationssystem *högst* en lösning $\mathbf{x}$.
Antag nämligen att både $\mathbf{x}$ och $\mathbf{y}$ är lösningar till (3.3) och sätt $\mathbf{z} = \mathbf{x} - \mathbf{y}$.
Då är $\mathbf{A}\mathbf{z} = \mathbf{A}(\mathbf{x} - \mathbf{y}) = \mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{y} = \mathbf{b} - \mathbf{b} = \mathbf{0}$, vilket betyder att $\mathbf{z}$ är en lösning till det homogena systemet (3.2). Men eftersom kolonnerna i $\mathbf{A}$ är linjärt oberoende så är därmed $\mathbf{z} = \mathbf{0}$, vilket betyder att $\mathbf{x} = \mathbf{y}$.

Notera att det faktum att kolonnerna i $\mathbf{A}$ är linjärt oberoende inte garanterar att det finns någon lösning $\mathbf{x}$ till ekvationssystemet (3.3). Men *om* det finns någon lösning så är den unik.

Om kolonnerna $\mathbf{a}_1, \ldots, \mathbf{a}_n$ i $m \times n$ matrisen $\mathbf{A}$ är linjärt oberoende så måste $n \leq m$, ty om $n > m$ så har ekvationssystemet (3.2) fler obekanta än ekvationer, och sådana "liggande" homogena ekvationssystem har alltid oändligt många lösningar.
(Se exempelvis avsnitt 1.4 i Petermanns bok Linjär geometri och algebra.)

3.2 Uppspännande vektorer i $\mathbb{R}^m$

Låt som ovan $\mathbf{a}_1, \dots, \mathbf{a}_n$ vara n st givna vektorer i $\mathbb{R}^m$.
Dessa vektorer säges *spänna upp* $\mathbb{R}^m$ om varje vektor $\mathbf{b} \in \mathbb{R}^m$ kan skrivas som en linjärkombination av dessa vektorer, dvs om det för varje vektor $\mathbf{b} \in \mathbb{R}^m$ finns n st tal $x_1, \dots, x_n$ sådana att

$$\mathbf{a}_1 x_1 + \cdots + \mathbf{a}_n x_n = \mathbf{b}. \tag{3.4}$$

Låt $\mathbf{A}$ vara en $m \times n$ matris med de givna vektorerna $\mathbf{a}_1, \dots, \mathbf{a}_n$ till kolonner, dvs $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$, och betrakta följande linjära ekvationssystem i variabelvektorn $\mathbf{x} = (x_1, \dots, x_n)^\mathsf{T} \in \mathbb{R}^n$:

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{3.5}$$

Eftersom (3.5) bara är ett alternativt sätt att skriva (3.4) på, så kan ovanstående definition ekvivalent uttryckas så här:
Kolonnerna till matrisen $\mathbf{A}$ spänner upp $\mathbb{R}^m$ om (och endast om) det för varje val av högerledsvektor $\mathbf{b} \in \mathbb{R}^m$ gäller att ekvationssystemet (3.5) har minst en lösning.

Om kolonnerna $\mathbf{a}_1, \dots, \mathbf{a}_n$ i $m \times n$ matrisen $\mathbf{A}$ spänner upp $\mathbb{R}^m$ så måste $n \geq m$, ty om $n < m$ så har ekvationssystemet (3.5) fler ekvationer än obekanta, och sådana "stående" ekvationssystem saknar lösning för vissa högerledsvektorer $\mathbf{b}$.
(Se exempelvis avsnitt 1.4 i Petermanns bok Linjär geometri och algebra.)

3.3 Baser till $\mathbb{R}^m$

Låt som ovan $\mathbf{a}_1, \dots, \mathbf{a}_n$ vara n st givna vektorer i $\mathbb{R}^m$.
Dessa vektorer säges utgöra en *bas* till $\mathbb{R}^m$ om de är linjärt oberoende *och* spänner upp $\mathbb{R}^m$.
Man kan direkt konstatera att om $\mathbf{a}_1, \dots, \mathbf{a}_n$ utgör en bas till $\mathbb{R}^m$ så är $n = m$, ty enligt ovan gäller att om $\mathbf{a}_1, \dots, \mathbf{a}_n$ spänner upp $\mathbb{R}^m$ så måste $n \geq m$, och om $\mathbf{a}_1, \dots, \mathbf{a}_n$ är linjärt oberoende så måste $n \leq m$.
Varje bas till $\mathbb{R}^m$ består alltså av exakt m st basvektorer.
Fortsättningsvis i detta avsnitt förutsätter vi därför att $n = m$.

Låt $\mathbf{A}$ vara en $m \times m$ matris med de givna vektorerna $\mathbf{a}_1, \dots, \mathbf{a}_m$ till kolonner, dvs $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_m \end{bmatrix}$, låt $\mathbf{b} \in \mathbb{R}^m$ vara en given högerledsvektor och betrakta följande ekvationssystem i variabelvektorn $\mathbf{x} = (x_1, \dots, x_m)^\mathsf{T} \in \mathbb{R}^m$:

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{3.6}$$

Antag att kolonnerna i $\mathbf{A}$ utgör en bas till $\mathbb{R}^m$.
Då spänner de upp $\mathbb{R}^m$, vilket innebär att systemet (3.6) har *minst* en lösning $\mathbf{x}$.
Vidare är de linjärt oberoende, vilket innebär att systemet (3.6) har *högst* en lösning.
Slutsatsen blir att om kolonnerna i $\mathbf{A}$ utgör en bas till $\mathbb{R}^m$ så gäller, för *varje* val av högerledsvektor $\mathbf{b} \in \mathbb{R}^m$, att ekvationssystemet (3.6) har *exakt* en lösning $\mathbf{x} \in \mathbb{R}^m$.
Detta är i sin tur ekvivalent med att varje vektor $\mathbf{b} \in \mathbb{R}^m$ på ett unikt sätt kan skrivas som en linjärkombination av basvektorerna $\mathbf{a}_1, \dots, \mathbf{a}_m$, dvs

$$\mathbf{a}_1 x_1 + \cdots + \mathbf{a}_m x_m = \mathbf{b},$$

där de reella talen $x_1, \ldots, x_m$ är unikt bestämda av $\mathbf{a}_1, \ldots, \mathbf{a}_m$ och $\mathbf{b}$.
Dessa tal $x_1, \ldots, x_m$ brukar kallas *koordinaterna* för $\mathbf{b}$ i basen $\mathbf{a}_1, \ldots, \mathbf{a}_m$.

Omvänt gäller att om det kvadratiska systemet (3.6) för varje val av högerledsvektor $\mathbf{b} \in \mathbb{R}^m$ har exakt en lösning $\mathbf{x} \in \mathbb{R}^m$ så utgör kolonnerna i $\mathbf{A}$ en bas till $\mathbb{R}^m$, ty att kolonnerna spänner upp $\mathbb{R}^m$ följer av att (3.6) har lösning för varje $\mathbf{b} \in \mathbb{R}^m$, och att kolonnerna är linjärt oberoende följer av att lösningen är unik även för högerledsvektorn $\mathbf{b} = \mathbf{0}$, så att endast $\mathbf{x} = \mathbf{0}$ uppfyller $\mathbf{A}\mathbf{x} = \mathbf{0}$.

3.4 Viktigt samband mellan kolonner och rader för en kvadratisk matris

För en *kvadratisk* matris $\mathbf{A} \in \mathbb{R}^{m \times m}$ är följande sex egenskaper ekvivalenta
(om en av dem gäller så gäller alla):

1. Kolonnerna i $\mathbf{A}$ är linjärt oberoende.
2. Kolonnerna i $\mathbf{A}$ spänner upp $\mathbb{R}^m$.
3. Kolonnerna i $\mathbf{A}$ utgör en bas till $\mathbb{R}^m$.
4. Kolonnerna i $\mathbf{A}^\mathsf{T}$ är linjärt oberoende.
5. Kolonnerna i $\mathbf{A}^\mathsf{T}$ spänner upp $\mathbb{R}^m$.
6. Kolonnerna i $\mathbf{A}^\mathsf{T}$ utgör en bas till $\mathbb{R}^m$.

Här ska bevisas att $\mathbf{1} \Rightarrow \mathbf{2} \Rightarrow \mathbf{4} \Rightarrow \mathbf{5} \Rightarrow \mathbf{1}$, vilket medför att $\mathbf{1} \Leftrightarrow \mathbf{2} \Leftrightarrow \mathbf{4} \Leftrightarrow \mathbf{5}$.
Därefter följer från definitionen av en bas att var och en av dessa fyra egenskaper är ekvivalent med var och en av de båda egenskaperna **3** och **6**.

$\mathbf{1} \Rightarrow \mathbf{2}$:
Antag att kolonnerna $\mathbf{a}_1, \ldots, \mathbf{a}_m$ i $\mathbf{A}$ är linjärt oberoende, och tag en godtycklig vektor $\mathbf{b} \in \mathbb{R}^m$. Vi ska visa att $\mathbf{b}$ är en linjärkombination av dessa kolonner.
Betrakta vektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}$. Eftersom detta är $m+1$ st vektorer i $\mathbb{R}^m$ så är de linjärt beroende. Alltså finns det $m+1$ st tal $x_1, \ldots, x_m, z$, som inte alla är 0, sådana att $\mathbf{a}_1 x_1 + \cdots + \mathbf{a}_n x_n + \mathbf{b} z = \mathbf{0}$.
Om här $z = 0$ så måste även alla $x_j = 0$ eftersom $\mathbf{a}_1, \ldots, \mathbf{a}_m$ är linjärt oberoende, och detta strider mot att minst ett av talen $x_1, \ldots, x_m, z$ är $\neq 0$.
Alltså är $z \neq 0$, varvid division med z ger att $\mathbf{b} = \mathbf{a}_1(-x_1/z) + \cdots + \mathbf{a}_n(-x_n/z)$, vilket visar att $\mathbf{b}$ är en linjärkombination av kolonnerna $\mathbf{a}_1, \ldots, \mathbf{a}_m$.

$\mathbf{2} \Rightarrow \mathbf{4}$:
Antag att kolonnerna i $\mathbf{A}$ spänner upp $\mathbb{R}^m$, och tag en godtycklig vektor $\mathbf{u} \in \mathbb{R}^m$ som uppfyller att $\mathbf{u} \neq \mathbf{0}$. Vi ska visa att $\mathbf{A}^\mathsf{T}\mathbf{u} \neq \mathbf{0}$. (Detta visar att den enda lösningen till $\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0}$ är $\mathbf{y} = \mathbf{0}$, vilket är just definitionen av att kolonnerna i $\mathbf{A}^\mathsf{T}$ är linjärt oberoende.)
Eftersom kolonnerna i $\mathbf{A}$ spänner upp $\mathbb{R}^m$ finns ett $\mathbf{x} \in \mathbb{R}^m$ sådant att $\mathbf{A}\mathbf{x} = \mathbf{u}$.
Men då uppfyller detta $\mathbf{x}$ även $\mathbf{u}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{u}^\mathsf{T}\mathbf{u}$ (som erhållits genom att multiplicera båda leden från vänster med $\mathbf{u}^\mathsf{T}$), dvs $\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{u} = |\mathbf{u}|^2 > 0$, vilket visar att $\mathbf{A}^\mathsf{T}\mathbf{u} \neq \mathbf{0}$.

$\mathbf{4} \Rightarrow \mathbf{5}$ bevisas genom att helt enkelt byta plats på $\mathbf{A}$ och $\mathbf{A}^\mathsf{T}$ i beviset av $\mathbf{1} \Rightarrow \mathbf{2}$ ovan.

$\mathbf{5} \Rightarrow \mathbf{1}$ bevisas genom att helt enkelt byta plats på $\mathbf{A}$ och $\mathbf{A}^\mathsf{T}$ i beviset av $\mathbf{2} \Rightarrow \mathbf{4}$ ovan.

3.5 Inversa matrisen $\mathbf{A}^{-1}$

Låt $\mathbf{A} \in \mathbb{R}^{m\times m}$ vara en kvadratisk matris och låt $\mathbf{I}$ beteckna $m\times m$ enhetsmatrisen.

Om det finns en matris $\mathbf{G} \in \mathbb{R}^{m\times m}$ sådan att $\mathbf{AG} = \mathbf{I}$ och $\mathbf{GA} = \mathbf{I}$ så säges $\mathbf{G}$ vara den *inversa* matrisen (eller kortare bara *inversen*) till $\mathbf{A}$. Beteckningen för detta är $\mathbf{G} = \mathbf{A}^{-1}$. Man har alltså att

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \;\text{ och }\; \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{3.7}$$

Det är lätt att visa att om $\mathbf{A}$ har en invers $\mathbf{A}^{-1}$ så är denna unik.
Om nämligen $\mathbf{AG} = \mathbf{I}$, $\mathbf{GA} = \mathbf{I}$, $\mathbf{AH} = \mathbf{I}$ och $\mathbf{HA} = \mathbf{I}$, så måste $\mathbf{G} = \mathbf{H}$, ty
$\mathbf{AG} = \mathbf{I} \Rightarrow \mathbf{H}(\mathbf{AG}) = \mathbf{H}\,\mathbf{I} \Rightarrow (\mathbf{HA})\mathbf{G} = \mathbf{H} \Rightarrow \mathbf{I}\,\mathbf{G} = \mathbf{H} \Rightarrow \mathbf{G} = \mathbf{H}$.

Här ska nu visas att $\mathbf{A}$ har en invers matris om och endast om $\mathbf{A}$ har någon (och därmed samtliga!) av de sex egenskaperna i föregående avsnitt.

Antag först att $\mathbf{A} \in \mathbb{R}^{m\times m}$ har de sex egenskaperna i föregående avsnitt.
Låt $\mathbf{e}_1, \ldots, \mathbf{e}_m$ beteckna de m st kolonnerna i enhetsmatrisen, så att $\mathbf{I} = \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_m \end{bmatrix}$.
Eftersom kolonnerna i $\mathbf{A}$ utgör en bas till $\mathbb{R}^m$ så finns det, för varje i, en unik vektor $\mathbf{y}_i \in \mathbb{R}^m$ sådan att $\mathbf{A}\mathbf{y}_i = \mathbf{e}_i$. Låt $\mathbf{Y}$ vara en $m\times m$ matris med dessa vektorer $\mathbf{y}_i$ till kolonner, dvs $\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_m \end{bmatrix}$. Då är alltså $\mathbf{AY} = \mathbf{I}$.
Eftersom även kolonnerna i $\mathbf{A}^\mathsf{T}$ utgör en bas till $\mathbb{R}^m$ så finns det, för varje i, en unik vektor $\mathbf{z}_i \in \mathbb{R}^m$ sådan att $\mathbf{A}^\mathsf{T}\mathbf{z}_i = \mathbf{e}_i$. Låt $\mathbf{Z}$ vara en $m\times m$ matris med dessa vektorer $\mathbf{z}_i$ till kolonner, dvs $\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 & \cdots & \mathbf{z}_m \end{bmatrix}$. Då är alltså $\mathbf{A}^\mathsf{T}\mathbf{Z} = \mathbf{I}$, vilket är ekvivalent med att $\mathbf{Z}^\mathsf{T}\mathbf{A} = \mathbf{I}$.
Det visar sig nu att $\mathbf{Z}^\mathsf{T} = \mathbf{Y}$, vilket kan visas på följande sätt:
Utgå från $\mathbf{AY} = \mathbf{I}$ och multiplicera bägge leden från vänster med $\mathbf{Z}^\mathsf{T}$.
Då erhålls att $\mathbf{Z}^\mathsf{T}\mathbf{AY} = \mathbf{Z}^\mathsf{T}$. Men eftersom $\mathbf{Z}^\mathsf{T}\mathbf{A} = \mathbf{I}$ så erhålls därmed att $\mathbf{Y} = \mathbf{Z}^\mathsf{T}$.
Matrisen $\mathbf{Y}$ uppfyller alltså både $\mathbf{AY} = \mathbf{I}$ och $\mathbf{YA} = \mathbf{I}$, vilket per definition innebär att $\mathbf{Y} = \mathbf{A}^{-1}$.

Antag nu omvänt att $\mathbf{A} \in \mathbb{R}^{m\times m}$ har en invers $\mathbf{A}^{-1}$ som alltså uppfyller (3.7).
Låt $\mathbf{b} \in \mathbb{R}^m$ vara en godtycklig högerledsvektor och betrakta ekvationssystemet

$$\mathbf{Ax} = \mathbf{b}. \tag{3.8}$$

Om vektorn $\mathbf{x} \in \mathbb{R}^m$ är en lösning till detta ekvationssystem så måste $\mathbf{x}$ även uppfylla $\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b}$, dvs $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ (eftersom $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$).
Men å andra sidan uppfyller detta unika $\mathbf{x}$ verkligen ekvationssystemet (3.8), ty $\mathbf{Ax} = \mathbf{A}\mathbf{A}^{-1}\mathbf{b} = \mathbf{b}$ (eftersom $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$).
Alltså har systemet (3.8), för varje högerledsvektor $\mathbf{b} \in \mathbb{R}^m$, en unik lösning $\mathbf{x}$.
Detta innebär att kolonnerna i matrisen $\mathbf{A}$ utgör en bas till $\mathbb{R}^m$.
Därmed har $\mathbf{A}$ en (och därmed samtliga) av de sex egenskaperna i föregående avsnitt.

Avslutningsvis kan man notera att matrisen $\mathbf{Z}$ ovan uppfyller både $\mathbf{A}^\mathsf{T}\mathbf{Z} = \mathbf{I}$ och $\mathbf{Z}\mathbf{A}^\mathsf{T} = \mathbf{I}$ (ty $\mathbf{Y}^\mathsf{T}\mathbf{A}^\mathsf{T} = \mathbf{I}$ och $\mathbf{Y}^\mathsf{T} = \mathbf{Z}$), vilket innebär att $\mathbf{Z}$ är inversen till $\mathbf{A}^\mathsf{T}$, dvs $\mathbf{Z} = (\mathbf{A}^\mathsf{T})^{-1}$.
Men $\mathbf{Z} = \mathbf{Y}^\mathsf{T} = (\mathbf{A}^{-1})^\mathsf{T}$, så vi får den viktiga räkneregeln

$$(\mathbf{A}^\mathsf{T})^{-1} = (\mathbf{A}^{-1})^\mathsf{T}. \tag{3.9}$$

3.6 För kvadratiska matriser gäller ekvivalensen $\mathbf{AG} = \mathbf{I} \;\Leftrightarrow\; \mathbf{GA} = \mathbf{I}$

Låt $\mathbf{A} \in I\!R^{m\times m}$ vara en given kvadratisk matris. Enligt ovan är den kvadratiska matrisen $\mathbf{G} \in I\!R^{m\times m}$ inversen till $\mathbf{A}$ om både $\mathbf{AG} = \mathbf{I}$ och $\mathbf{GA} = \mathbf{I}$.
Det visar sig dock att man bara behöver kräva ett av dessa bägge samband, ty om ett av dem är uppfyllt så är även det andra det!

Antag till att börja med att $\mathbf{GA} = \mathbf{I}$.
Då har $\mathbf{A}$ linjärt oberoende kolonner, ty vi har att $\mathbf{Ax} = \mathbf{0} \;\Rightarrow\; \mathbf{GAx} = \mathbf{G0} \;\Rightarrow$
$\Rightarrow\; \mathbf{I}\,\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{x} = \mathbf{0}$, vilket visar att den enda lösningen till $\mathbf{Ax} = \mathbf{0}$ är $\mathbf{x} = \mathbf{0}$.
Därmed har $\mathbf{A}$ en av de sex egenskaperna ovan, vilket medför att $\mathbf{A}$ har en invers $\mathbf{A}^{-1}$.
Multiplikation från höger med denna invers ger att $\mathbf{GA} = \mathbf{I} \;\Rightarrow\; \mathbf{GAA}^{-1} = \mathbf{IA}^{-1} \;\Rightarrow$
$\Rightarrow\; \mathbf{G} = \mathbf{A}^{-1}$, vilket visar att $\mathbf{G}$ är inversen till $\mathbf{A}$. Därmed gäller även $\mathbf{AG} = \mathbf{I}$.

Antag nu i stället att $\mathbf{AG} = \mathbf{I}$, vilket är ekvivalent med att $\mathbf{G}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}} = \mathbf{I}$.
Då har $\mathbf{A}^{\mathsf{T}}$ linjärt oberoende kolonner, ty vi har att $\mathbf{A}^{\mathsf{T}}\mathbf{u} = \mathbf{0} \;\Rightarrow\; \mathbf{G}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\mathbf{u} = \mathbf{G}^{\mathsf{T}}\mathbf{0} \;\Rightarrow$
$\Rightarrow\; \mathbf{I}\,\mathbf{u} = \mathbf{0} \;\Rightarrow\; \mathbf{u} = \mathbf{0}$, vilket visar att den enda lösningen till $\mathbf{A}^{\mathsf{T}}\mathbf{u} = \mathbf{0}$ är $\mathbf{u} = \mathbf{0}$.
Därmed har $\mathbf{A}$ en av de sex egenskaperna ovan, vilket medför att $\mathbf{A}$ har en invers $\mathbf{A}^{-1}$.
Multiplikation från vänster med denna invers ger att $\mathbf{AG} = \mathbf{I} \;\Rightarrow\; \mathbf{A}^{-1}\mathbf{AG} = \mathbf{A}^{-1}\,\mathbf{I} \;\Rightarrow$
$\Rightarrow\; \mathbf{G} = \mathbf{A}^{-1}$, vilket visar att $\mathbf{G}$ är inversen till $\mathbf{A}$. Därmed gäller även $\mathbf{GA} = \mathbf{I}$.

Vi kan nu komplettera listan av ekvivalenta egenskaper för en kvadratisk matris enligt följande:

För en *kvadratisk* matris $\mathbf{A} \in I\!R^{m\times m}$ är följande nio egenskaper ekvivalenta
(om en av dem gäller så gäller alla):

1. Kolonnerna i $\mathbf{A}$ är linjärt oberoende.
2. Kolonnerna i $\mathbf{A}$ spänner upp $I\!R^m$.
3. Kolonnerna i $\mathbf{A}$ utgör en bas till $I\!R^m$.
4. Kolonnerna i $\mathbf{A}^{\mathsf{T}}$ är linjärt oberoende.
5. Kolonnerna i $\mathbf{A}^{\mathsf{T}}$ spänner upp $I\!R^m$.
6. Kolonnerna i $\mathbf{A}^{\mathsf{T}}$ utgör en bas till $I\!R^m$.
7. Det finns en matris $\mathbf{G} \in I\!R^{m\times m}$ sådan att $\mathbf{AG} = \mathbf{I}$.
8. Det finns en matris $\mathbf{G} \in I\!R^{m\times m}$ sådan att $\mathbf{GA} = \mathbf{I}$.
9. $\mathbf{A}$ är inverterbar, dvs har en invers $\mathbf{A}^{-1}$.

Det finns en tionde egenskap som också är ekvivalent med var och en av ovanstående nio.
Vi har ännu inte tagit upp den, men nämner den här för fullständighets skull.

10. Determinanten för matrisen $\mathbf{A}$ är $\neq 0$.

En kvadratisk matris $\mathbf{A}$ som har en av dessa egenskaper (och därmed samtliga) säges vara *icke-singulär*.