نگاشت‌های خطی (۱)

وقتی $m \times n$ عدد حقیقی a_{ij}، $i = ۱, \ldots, m$، $j = ۱, \ldots, n$، را در یک قالب m در n، با m ردیف و n ستون به شکل زیر بنویسیم، یک ماتریس $m \times n$ (حقیقی) تعریف می‌شود:

$$\begin{bmatrix} a_{۱۱} & \ldots & a_{۱n} \\ \vdots & & \vdots \\ a_{m۱} & & a_{mn} \end{bmatrix}$$

اعداد مندرج در ماتریس را درایه‌های ماتریس می‌نامند. مقصود از a_{ij} درایه‌ای است که در ردیف i ام و ستون j ام قرار دارد. گاهی اوقات A را به صورت $[a_{ij}]$، یا به طور کاملتر $[a_{ij}]_{i=۱,\ldots,m\ j=۱,\ldots,n}$ می‌نویسیم.

برای ماتریس‌های هم اندازه می‌توان عمل جمع، و عمل ضرب در اعداد حقیقی، همانند عملیات مشابه برای n تایی‌ها (بردارها) را تعریف کرد. اگر $A = [a_{ij}]$ و $B = [b_{ij}]$ دو ماتریس $m \times n$ باشند، مجموع دو ماتریس، $A + B = [c_{ij}]$ با فرمول

$$c_{ij} = a_{ij} + b_{ij} \tag{۱}$$

تعریف می‌شود. تحقیق صحت خواص زیر از تعریف سرراست است:

(۱۱–۱) خواص ابتدایی جمع ماتریس‌ها

(۱۱–۱–۱) تعویض‌پذیری (جابجایی). اگر A و B دو ماتریس $m \times n$ باشند، داریم

$A + B = B + A$.

(۱۱—۱—۲) شرکت‌پذیری. اگر A، B و C ماتریس‌های $m \times n$ باشند، داریم $A + (B + C) = (A + B) + C$.

(۱۱—۱—۳) عنصر بی‌اثر. ماتریس صفر $m \times n$ که همهٔ درایه‌های آن صفر است و به O نمایش می‌دهیم (یگانه ماتریس $m \times n$) دارای این ویژگی است که برای هر ماتریس $m \times n$، A، داریم $A + O = O + A = A$.

(۱۱—۱—۴) عنصر قرینه. برای $A = [a_{ij}]$، ماتریس $-A = [b_{ij}]$ که با $b_{ij} = -a_{ij}$ تعریف می‌شود (یگانه ماتریس) دارای این ویژگی است که $A + (-A) = (-A) + A = O$.

حال اگر $A = [a_{ij}]$ یک ماتریس $m \times n$ باشد و $r \in \mathbb{R}$، ماتریس $rA = [b_{ij}]$ با فرمول $b_{ij} = ra_{ij}$ تعریف می‌شود، یعنی همهٔ درایه‌های A در عدد r ضرب می‌شوند. تحقیق خواص زیر نیز سرراست است:

(۱۱—۲) خواص ابتدایی ضرب اعداد در ماتریس‌ها

(۱۱—۲—۱) $1 \cdot A = A$ برای هر ماتریس A.

(۱۱—۲—۲) $(rs)A = r(sA)$ برای هر ماتریس A و هر زوج عدد حقیقی r, s.

(۱۱—۲—۳) $(r + s)A = rA + sA$ برای هر ماتریس A و هر زوج عدد حقیقی r, s.

(۱۱—۲—۴) $r(A + B) = rA + rB$ برای هر عدد حقیقی r و هر دو ماتریس هم اندازهٔ A و B.

علاوه بر عملیات بالا، مفهوم ضرب ماتریس‌ها نیز مابین ماتریس‌های اندازه‌های مناسب که توضیح

داده خواهد شد، تعریف می‌شود. اگر $A=[a_{ij}]$ یک ماتریس $m\times n$ و $B=[b_{ij}]$ یک ماتریس $n\times p$ باشد، ماتریس حاصل‌ضرب، $AB=[c_{ij}]$، به صورت زیر تعریف می‌شود:

$$c_{ij}=a_{i۱}b_{۱j}+a_{i۲}b_{۲j}+\cdots+a_{in}b_{nj} \qquad (۲)$$

توجه کنید که لازمهٔ تعریف حاصل‌ضرب این است که تعداد ستون‌های A برابر تعداد ردیف‌های B باشد، بنابراین عبارت (۲) معنی دارد و در واقع می‌توان c_{ij} را حاصل‌ضرب داخلی ردیف i ام ماتریس A به عنوان n تایی $(a_{i۱},\ldots,a_{in})$ با ستون j ام ماتریس B، به عنوان n تایی $(b_{j۱},\ldots,b_{nj})$ تلقی کرد. انگیزهٔ این تعریف را در جلسهٔ آینده بررسی خواهیم کرد.

(۱۱—۳) خواص ابتدایی ضرب ماتریس‌ها

(۱۱—۳—۱) شرکت‌پذیری. A، B و C ماتریس‌های دارای اندازهٔ به ترتیب $m\times n$، $n\times p$ و $p\times q$ هستند. در این صورت:

$$(AB)C=A(BC)$$

(۱۱—۳—۲) قانون پخشی. داریم

$$(A+B)C=(AC)+(AB) \quad , \quad A(B+C)=(AB)+(AC)$$

مشروط بر این که اندازهٔ ماتریس‌های فوق برای عملیات ذکر شده مناسب باشد.

(۱۱—۳—۳) ماتریس واحد $n\times n$، $I_n=[\delta_{ij}]$ بدین صورت تعریف می‌شود که:

$$\delta_{ij}=\begin{cases} ۱ & i=j \\ ۰ & i\neq j \end{cases}$$

برای هر ماتریس $m \times n$، A، و هر ماتریس $n \times p$، B، داریم:

$$A \cdot I_n = A \quad , \quad I_n \cdot B = B$$

اثبات (۱۱—۳—۱) را ارائه می‌کنیم، دوتای دیگر ساده‌ترند و تحقیق آنها به خواننده واگذار می‌شود. برای (۱۱—۳—۱) بنویسید $A = [a_{ij}]$، $B = [b_{ij}]$، $C = [c_{ij}]$، $AB = [d_{ij}]$، $BC = [e_{ij}]$، $(AB)C = [x_{ij}]$ و $A(BC) = [y_{ij}]$. طبق تعریف حاصل‌ضرب ماتریس‌ها عمل می‌کنیم:

$$\begin{aligned} x_{ij} = &= \sum_{i=1}^{p} d_{ik}c_{kj} \\ &= \sum_{k=1}^{p}(\sum_{l=1}^{n} a_{il}b_{lk})c_{kj} \\ &= \sum_{k=1}^{p}\sum_{l=1}^{n} a_{il}b_{lk}c_{kj} \\ &= \sum_{l=1}^{n} a_{il}(\sum_{k=1}^{p} b_{lk}c_{kj}) \\ &= \sum_{l=1}^{n} a_{il}e_{lj} \\ &= y_{ij} \end{aligned}$$

و حکم به اثبات می‌رسد.

در اینجا قابل ذکر است که اگر AB تعریف شده باشد، لزومی ندارد BA نیز تعریف شدنی باشد زیرا که تعریف AB مستلزم این است که تعداد ستون‌های A برابر تعداد ردیف‌های B باشد، و این حکمی در مورد برابری تعداد ستون‌های B با تعداد ردیف‌های A نمی‌کند. حتی اگر AB و BA هر دو تعریف‌شدنی باشند، لزومی ندارد هر دو یک اندازه باشند، مثلاً اگر A یک ماتریس $n \times ۱$ و B یک

ماتریس $n \times ۱$ باشند، $n > ۱$،

$$A = [a_۱ \ \ldots \ a_n] \quad , \quad B = \begin{bmatrix} b_۱ \\ \vdots \\ b_n \end{bmatrix}$$

AB یک ماتریس ۱×۱، با تک درایهٔ $\sum_{i=۱}^{n} a_i b_i$ می‌شود، در حالی که BA یک ماتریس $n \times n$ است که درایهٔ ردیف i ام و ستون j ام آن عبارت است از $b_j a_i$. حتی اگر A و B هر دو $n \times n$، با $n \geq ۲$، باشند، لزومی بر تساوی AB و BA نیست، مثلاً:

$$\begin{bmatrix} ۱ & ۱ \\ ۰ & ۰ \end{bmatrix}\begin{bmatrix} ۰ & ۰ \\ ۱ & ۱ \end{bmatrix} = \begin{bmatrix} ۱ & ۱ \\ ۰ & ۰ \end{bmatrix} \quad , \quad \begin{bmatrix} ۰ & ۰ \\ ۱ & ۱ \end{bmatrix}\begin{bmatrix} ۱ & ۱ \\ ۰ & ۰ \end{bmatrix} = \begin{bmatrix} ۰ & ۰ \\ ۱ & ۱ \end{bmatrix}$$

البته در مورد ماتریس I_n و هر ماتریس $n \times n$، A، از ۱۱—۳—۳ نتیجه می‌شود که $AI_n = I_n A$ $(A =)$.

توجه کنید که برای یک nـ تایی مرتب از اعداد، مثلاً $x = (x_۱, \ldots, x_n)$، اکنون سه نمایش ممکن داریم، یکی به صورت بالا که معمولاً تجسم برداری دارد، دیگر به صورت یک ماتریس $۱ \times n$، $[x_۱ \ \ldots \ x_n]$ و بالاخره به صورت یک ماتریس $n \times ۱$، $\begin{bmatrix} x_۱ \\ \vdots \\ x_n \end{bmatrix}$. اغلب لازم است که بین این سه نمایش تمایز قائل شویم تا ابهامی در عملیات حاصل نشود. برای این منظور، هرگاه n تایی $x = (x_۱, \ldots, x_n)$ به صورت یک ماتریس $n \times ۱$ نوشته شود، آن را به $|x\rangle$ نمایش می‌دهیم، و هرگاه به صورت یک ماتریس $۱ \times n$ نمایش داده شود، آن را به $\langle x|$ نمایش می‌دهیم.

با این مقدمه در مورد ماتریس‌ها، به موضوع اصلی درس که "نگاشت‌های خطی" است می‌پردازیم. کلمات "تابع"، "نگاشت" و "تبدیل" در اینجا به صورت کاملاً مترادف به‌کار خواهند رفت هر چند که به لحاظ تاریخی گاهی تصاویر ذهنی متفاوتی از این کلمات القاء می‌شود. اگر S و T دو مجموعه باشند، مقصود از یک تابع (= نگاشت، = تبدیل) از S به T، قانونی f است که به هر عنصر s از S، عنصر مشخصی $f(s)$ از T نسبت می‌دهد. می‌نویسیم $f : S \longrightarrow T$. در گذشته کلمهٔ "تابع" معمولاً وقتی به‌کار می‌رفته است که T برابر $\mathbb{R}$ یا $\mathbb{C}$ باشد، "تبدیل" در حالت $S = T$، یعنی تابع‌های از یک مجموعه

به خود آن، و "نگاشت" برای تابع‌های $\mathbb{R}^n \longrightarrow \mathbb{R}^m$، m, n بزرگتر از ۱، به‌کار می‌رفته است. ما چنین تمایزهایی قایل نمی‌شویم. برای تابع $f : S \longrightarrow T$، مجموعهٔ S را دامنه (= قلمرو)، و مجموعهٔ T را بُرد می‌نامیم. مجموعه $\{f(s) \mid s \in S\}$ که یک زیرمجموعهٔ T است، تصویر f، یا تصویر S تحت f خوانده می‌شود. در واقع برای هر زیرمجموعهٔ E از S، تصویر E تحت f، که به $f(E)$ نمایش داده می‌شود به صورت:

$$f(E) = \{f(s) \mid s \in E\}$$

تعریف می‌شود. تابع f را یک به یک می‌نامیم در صورتی که $f(s_1) = f(s_2)$ همواره نتیجه دهد $s_1 = s_2$، و f پوشا می‌نامیم اگر $f(S) = T$. برای تابع‌های یک به یک و پوشا، $f : S \longrightarrow T$، تابعی به نام وارون f (یا معکوس f) به صورت زیر تعریف می‌شود که به $f^{-1} : T \longrightarrow S$ نمایش می‌دهیم. هرگاه $t \in T$، چون f پوشا است، عنصری s از S وجود دارد که $f(s) = t$. به علاوه چون f یک به یک فرض شده است، چنین عنصر s به طور منحصر به فرد تعیین می‌شود، پس f^{-1} به صورت قانونی بی‌ابهام $f^{-1}(t) = s$ تعریف شدنی است و داریم:

$$f \circ f^{-1} = 1_T \quad , \quad f^{-1} \circ f = 1_S$$

که مقصود از 1_X تابع همانی مجموعهٔ X است، $1_X(x) = x$ برای هر $x \in X$.

نماد $f^{-1}(t)$ حتی وقتی f^{-1} تعریف شدنی نباشد نیز به‌کار می‌رود. مقصود از $f^{-1}(t)$ مجموعهٔ نقاط S است که تحت f به t فرستاده (= نگاشته) می‌شوند، یعنی:

$$f^{-1}(t) = \{s \in S \mid f(s) = t\}$$

$f^{-1}(t)$ را گاهی مجموعهٔ تراز منسوب به t می‌خوانیم. به طور کلی، اگر Y زیرمجموعه‌ای از T باشد، زیرمجموعهٔ $f^{-1}(Y)$ (تصویر وارون Y) یک زیرمجموعهٔ S است که بدین صورت تعریف می‌شود:

$$f^{-1}(Y) = \{s \in S \mid f(s) \in Y\}$$

موضوع اصلی درس ما بررسی تابع‌های $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ (یا به طور کلی‌تر، تابع‌های $f:S \longrightarrow \mathbb{R}^m$، $S \subset \mathbb{R}^n$) است. بدین ترتیب چنین تابع f به هر n تایی مرتب از اعداد حقیقی در دامنه، یک m تایی مرتب از اعداد حقیقی نسبت می‌دهد:

$$f(x_1,\ldots,x_n) = (y_1,\ldots,y_m) \tag{۳}$$

هر y_i وابسته به $(x_1,\ldots,x_n)$ است، پس در واقع ارائه f همانند ارائه m تابع $f_1,\ldots,f_m$، هر یک از $\mathbb{R}^n$ به $\mathbb{R}$ است، $f_i(x_1,\ldots,x_n) = y_i$، $i=1,\ldots,m$. (۳) را به صورت زیر نیز می‌نویسیم:

$$\begin{cases} f_1(x_1,\ldots,x_n) = y_1 \\ \quad\ldots \\ f_m(x_1,\ldots,x_n) = y_m \end{cases} \tag{۴}$$

گاهی می‌نویسیم $f=(f_1,\ldots,f_m)$ و هر f_i را یک مؤلفه f می‌نامیم.

تابع $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ در (۳) یا (۴) را خطی می‌نامیم در صورتی که هر $f_i(x_1,\ldots,x_n)$ یک عبارت همگن درجه اول نسبت به $x_1,\ldots,x_n$ باشد، یعنی:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \quad\ldots \quad \ldots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases} \tag{۵}$$

مقصود از "همگن" این است که همهٔ جملات سمت راست از یک درجه‌اند، در اینجا همه از درجه ۱، بالاخص جملهٔ ثابت وجود ندارد. اگر هر عبارت سمت راست را فقط درجه ۱ در نظر بگیریم و محدودیت همگن را حذف کنیم، یک عدد ثابت نیز ممکن است به هر جمله افزوده شود. این گونه تابع‌ها را مستوی می‌نامیم:

$$\begin{cases} y_1 = a_{10} + a_{11}x_1 + \cdots + a_{1n}x_n \\ \quad\ldots \quad \ldots \\ y_m = a_{m0} + a_{m1}x_1 + \cdots + a_{mn}x_n \end{cases} \tag{۶}$$

مقدار تابع مستوی (٦) از انتقال مقدار تابع (۵) با mـ تایی ثابت $(a_{10}, \dots, a_{m0})$ به‌دست می‌آید و نیازی به بررسی جداگانهٔ تابع‌های مستوی نیست.

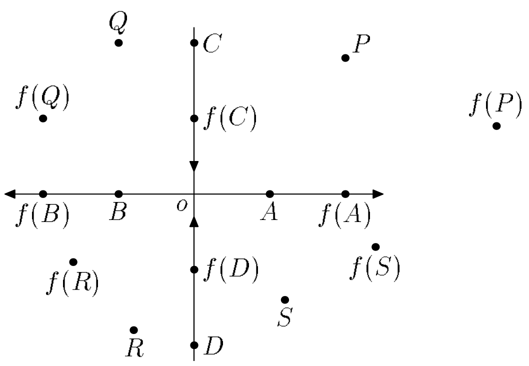
(۱۱—۴) چند مثال

(۱۱—۴—۱) هر تابع خطی $f : \mathbb{R} \longrightarrow \mathbb{R}$ به شکل $f(x) = mx$ است که در آن $m \in \mathbb{R}$ یک عدد حقیقی داده شده است. تابع‌های مستوی $g : \mathbb{R} \longrightarrow \mathbb{R}$ به شکل $g(x) = mx + b$، m و b اعداد حقیقی داده شده، می‌باشند.

(۱۱—۴—۲) $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ را به صورت زیر تعریف می‌کنیم

$$f(x_1, x_2) = (2x_1, \frac{1}{2}x_2)$$

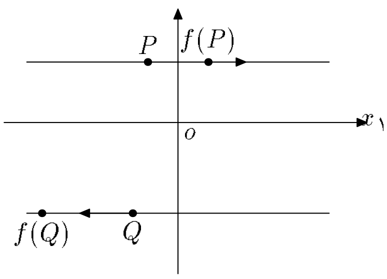
هر دو عبارت $2x_1$ و $\frac{1}{2}x_2$ درجه ۱ همگن نسبت به (x_1, x_2) هستند، پس f خطی است. عملکرد این تابع بدین صورت است که فاصله هر نقطه را از محور x_2 به دو برابر افزایش می‌دهد و فاصلهٔ آن از محور x_1 را نصف می‌کند. محور x_1 با ضریب تجانس ۲ به خود آن نگاشته می‌شود (انبساط با ضریب ۲)، و محور x_2 با ضریب تجانس $\frac{1}{2}$ روی خود منقبض می‌شود. (شکل ۱)



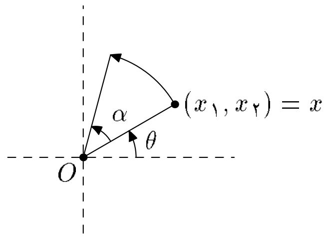
(۱۱–۴–۲) $f : \mathbb{R}^{۲} \longrightarrow \mathbb{R}^{۲}$ را به صورت زیر تعریف می‌کنیم:

$$f(x_{۱}, x_{۲}) = (x_{۱} + x_{۲}, x_{۲})$$

f خطی است. توجه کنید که هر خط راست افقی $x_{۲} = c$ به خود آن نگاشته می‌شد ولی هر نقطه روی این خط به اندازه مقدار c انتقال افقی می‌یابد. بدین ترتیب خطوط راست افقی بالای محور $x_{۱}$ به طرف راست و خطوط افقی پایین محور $x_{۱}$ به سمت چپ روی خود می‌لغزند. نقاط محور $x_{۱}$ سر جای خود ثابت می‌مانند.



(۱۱–۴–۳) (دوران حول $\underline{o}$ در $\mathbb{R}^{۲}$) دوران حول $\underline{o}$ در $\mathbb{R}^{۲}$ با زاویهٔ α را در نظر بگیرید. اگر $x = (x_{۱}, x_{۲})$ نقطه‌ای $\neq \underline{o}$ در صفحه باشد، نمایش آن به صورت قطبی در نظر بگیرید: $x_{۱} = |x| \cos\theta$ ،$x_{۲} = |x| \sin\theta$.



اگر دوران زاویهٔ α حول $\underline{o}$ را به $f : \mathbb{R}^{۲} \longrightarrow \mathbb{R}^{۲}$ نمایش دهیم، نمایش قطبی $f(x_{۱}, x_{۲})$ می‌شود:

$$\begin{aligned} f(x_{۱}, x_{۲}) &= (|x|\cos(\theta+\alpha), |x|\sin(\theta+\alpha)) \\ &= (|x|\cos\theta\cos\alpha - |x|\sin\theta\sin\alpha, |x|\sin\theta\cos\alpha + |x|\cos\theta\sin\alpha) \\ &= ((\cos\alpha)x_{۱} - (\sin\alpha)x_{۲}, (\sin\alpha)x_{۱} + (\cos\alpha)x_{۲}) \end{aligned}$$

هر یک از دو مؤلفهٔ f یک تابع درجه یک همگن نسبت به $(x_۱, x_۲)$ است، پس f خطی است.

(۱۱–۴–۴) (تصویر قائم روی یک زیرفضای مختصاتی) فرض کنید $n > m$ و $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ را به صورت زیر تعریف کنید

$$f(x_۱, \ldots, x_n) = (x_۱, \ldots, x_m)$$

این تابع خطی است و عملکرد آن این است که نقطهٔ $(x_۱, \ldots, x_n)$ را به $(x_۱, \ldots, x_m)$، متشکل از m مؤلفهٔ اول آن، می‌نگارد. مثلاً برای $n = ۳$ و $m = ۲$، $f(x, y, z) = (x, y)$ تصویر قائم (x, y, z) روی صفحهٔ xy است.

(۱۱–۴–۵) $f : \mathbb{R}^۲ \longrightarrow \mathbb{R}^۴$ را به صورت زیر تعریف کنید:

$$f(x_۱, x_۲) = (x_۱, x_۱ - x_۲, ۲x_۱ + x_۲, -x_۲)$$

چون هر مؤلفه f یک عبارت درجهٔ یک همگن نسبت به $(x_۱, x_۲)$ است، این تابع خطی است.

(۱۱–۴–٦) $f : \mathbb{R}^۳ \longrightarrow \mathbb{R}^۲$ به صورت زیر تعریف می‌شود:

$$f(x_۱, x_۲, x_۳) = (x_۱ - x_۲ + x_۳, x_۱ x_۲)$$

این تابع خطی نسبت زیرا مؤلفه دوم f، یعنی $x_۱ x_۲$ از درجهٔ یک نیست.

اکنون ارتباط میان دو بحث این جلسه را بیان می‌کنیم. توجه کنید که می‌توان (۵) را به صورت زیر نوشت:

$$\begin{bmatrix} y_۱ \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{۱۱} & \ldots & a_{۱n} \\ \vdots & & \vdots \\ a_{m۱} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_۱ \\ \vdots \\ x_n \end{bmatrix} \tag{۷}$$

یا به طور موجز:

$$|y\rangle = A|x\rangle \tag{۸}$$

یعنی محاسبهٔ مقدار یک تابع خطی بدین صورت حاصل می‌شود که ضرایب a_{ij} در (۵) را در یک ماتریس $m \times n$ قرار می‌دهیم و حاصل‌ضرب این ماتریس را با ستون $\langle x|$ محاسبه می‌کنیم. بدین ترتیب یک تناظر یک به یک میان تابع‌های خطی $\mathbb{R}^m \longrightarrow \mathbb{R}^n$ و ماتریس‌های $m \times n$ ایجاد می‌شود.

برای مثال‌های خطی بالا، ماتریس‌های مربوط به ترتیب از بالا به پایین عبارتند از:

$$[m]$$

$$\begin{bmatrix} ۲ & ۰ \\ ۰ & \frac{۱}{۲} \end{bmatrix}$$

$$\begin{bmatrix} ۱ & ۱ \\ ۰ & ۱ \end{bmatrix}$$

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

$$\begin{bmatrix} ۱ & & O & \\ & \ddots & & O \\ O & & ۱ & \end{bmatrix}$$

$$\begin{bmatrix} ۱ & ۰ \\ ۱ & -۱ \\ ۲ & ۱ \\ ۰ & -۱ \end{bmatrix}$$

با نگاه کردن به ماتریس متناظر به یک تابع خطی می‌توان عمل تابع خطی بر اعضای پایهٔ متداول، یعنی $e_۱, \ldots, e_n$، را فوراً دریافت. توجه کنید که اگر ماتریس $A = [a_{ij}]$ از سمت چپ در ستون $\langle e_j|$ ضرب شود، ستون jــ ام ماتریس $\begin{bmatrix} a_{۱j} \\ \vdots \\ a_{mj} \end{bmatrix}$ به‌دست می‌آید. بدین ترتیب، برای نوشتن ماتریس مربوط به یک تابع خطی f، کافی است $f(e_۱)$، ... ، $f(e_n)$ را محاسبه کنیم و نتایج حاصل را به ترتیب در ستون‌های اول تا n ام درج کنیم.