MATH 532: Linear Algebra

Chapter 6: Determinants

Greg Fasshauer

Department of Applied Mathematics
Illinois Institute of Technology

Spring 2015

Determinants

We will concentrate on a few — not so well-known, but useful — facts about determinants.

Theorem

If A *and* D *are square matrices (not necessarily of the same size) such that* A^{-1} *exists, then*

$$\det \begin{pmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{pmatrix} = \det(\mathsf{A}) \det(\mathsf{D} - \mathsf{C}\mathsf{A}^{-1}\mathsf{B}),$$

where $\mathsf{S} = \mathsf{D} - \mathsf{C}\mathsf{A}^{-1}\mathsf{B}$ *is the Schur complement of* A *(cf. HW 2).*

Determinants

We will concentrate on a few — not so well-known, but useful — facts about determinants.

Theorem

If A *and* D *are square matrices (not necessarily of the same size) such that* A^{-1} *exists, then*

$$\det \begin{pmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{pmatrix} = \det(\mathsf{A}) \det(\mathsf{D} - \mathsf{C}\mathsf{A}^{-1}\mathsf{B}),$$

where $\mathsf{S} = \mathsf{D} - \mathsf{C}\mathsf{A}^{-1}\mathsf{B}$ *is the Schur complement of* A *(cf. HW 2).*

Remark

An analogous formula exists if D *is invertible instead of* A*.*

Proof

Idea: Use block LU factorization because we know that:

Proof

Idea: Use block LU factorization because we know that:

- Determinant of a product is product of determinants, i.e.,

$$\det(A_1 A_2) = \det(A_1)\det(A_2).$$

Proof

Idea: Use block LU factorization because we know that:

- Determinant of a product is product of determinants, i.e.,

$$\det(A_1 A_2) = \det(A_1)\det(A_2).$$

- Determinant of a diagonal matrix is product of diagonal elements. This is also true for block matrices if we use determinants, i.e.,

$$\det\begin{pmatrix} U_1 & U_2 \\ O & U_3 \end{pmatrix} = \det(U_1)\det(U_3).$$

Proof (cont.)

Similarly to the midterm exam, we get

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & O \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ O & D - CA^{-1}B \end{pmatrix}.$$

Proof (cont.)

Similarly to the midterm exam, we get

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & O \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ O & D - CA^{-1}B \end{pmatrix}.$$

But then

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \underbrace{\det(I)}_{=1} \underbrace{\det(I)}_{=1} \det(A) \det(D - CA^{-1}B).$$

□

Example

in the context of kriging (cf. HW 2), the (scaled) kriging variance

$$K(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}(\boldsymbol{x})^T \mathsf{K}^{-1} \boldsymbol{k}(\boldsymbol{x})$$

can be computed as a ratio of determinants (see the multipage proof in [Sch05, FWL04]).

Example

in the context of kriging (cf. HW 2), the (scaled) kriging variance

$$K(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}(\boldsymbol{x})^T \mathsf{K}^{-1} \boldsymbol{k}(\boldsymbol{x})$$

can be computed as a ratio of determinants (see the multipage proof in [Sch05, FWL04]).

Simple two line proof:
Let $\mathsf{A} = \mathsf{K} \in \mathbb{R}^{n \times n}$, $\mathsf{B} = \boldsymbol{k}(\boldsymbol{x}) \in \mathbb{R}^{n \times 1}$, $\mathsf{C} = \boldsymbol{k}(\boldsymbol{x})^T \in \mathbb{R}^{1 \times n}$, $\mathsf{D} = K(\boldsymbol{x}, \boldsymbol{x}) \in \mathbb{R}$, i.e.,

$$\det \begin{pmatrix} \mathsf{K} & \boldsymbol{k}(\boldsymbol{x}) \\ \boldsymbol{k}(\boldsymbol{x})^T & K(\boldsymbol{x}, \boldsymbol{x}) \end{pmatrix} = \det(\mathsf{K}) \det\left(K(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}(\boldsymbol{x})^T \mathsf{K}^{-1} \boldsymbol{k}(\boldsymbol{x})\right)$$

Example

in the context of kriging (cf. HW 2), the (scaled) kriging variance

$$K(\boldsymbol{x},\boldsymbol{x}) - \boldsymbol{k}(\boldsymbol{x})^T \mathsf{K}^{-1} \boldsymbol{k}(\boldsymbol{x})$$

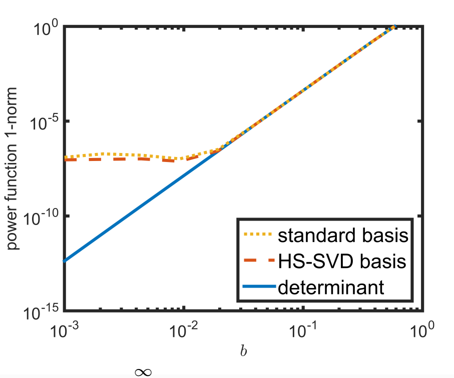
can be computed as a ratio of determinants (see the multipage proof in [Sch05, FWL04]).

Simple two line proof:
Let $\mathsf{A} = \mathsf{K} \in \mathbb{R}^{n\times n}$, $\mathsf{B} = \boldsymbol{k}(\boldsymbol{x}) \in \mathbb{R}^{n\times 1}$, $\mathsf{C} = \boldsymbol{k}(\boldsymbol{x})^T \in \mathbb{R}^{1\times n}$, $\mathsf{D} = K(\boldsymbol{x},\boldsymbol{x}) \in \mathbb{R}$, i.e.,

$$\det \begin{pmatrix} \mathsf{K} & \boldsymbol{k}(\boldsymbol{x}) \\ \boldsymbol{k}(\boldsymbol{x})^T & K(\boldsymbol{x},\boldsymbol{x}) \end{pmatrix} = \det(\mathsf{K}) \det\left(K(\boldsymbol{x},\boldsymbol{x}) - \boldsymbol{k}(\boldsymbol{x})^T \mathsf{K}^{-1} \boldsymbol{k}(\boldsymbol{x})\right)$$

$$\iff \quad K(\boldsymbol{x},\boldsymbol{x}) - \boldsymbol{k}(\boldsymbol{x})^T \mathsf{K}^{-1} \boldsymbol{k}(\boldsymbol{x}) = \frac{\det \begin{pmatrix} \mathsf{K} & \boldsymbol{k}(\boldsymbol{x}) \\ \boldsymbol{k}(\boldsymbol{x})^T & K(\boldsymbol{x},\boldsymbol{x}) \end{pmatrix}}{\det(\mathsf{K})}.$$

Computing the kriging variance — Example [FM15]



Analytic Chebyshev kernel $K(x,z) = \sum_{n=0}^{\infty} \lambda_n \varphi_n(x)\varphi_n(z)$ on 11 Chebyshev points in $[-1,1]$

$$\lambda_0 = \frac{1}{2}, \quad \lambda_n = \frac{(1-b)b^n}{2b}, \quad \varphi_n(x) = \sqrt{2-\delta_{n0}}\,T_n(x),$$

$$K(x,z) = \frac{1}{2} + (1-b)\frac{b(1-b^2) - 2b(x^2+z^2) + (1+3b^2)xz}{(1-b^2)^2 + 4b\left(b(x^2+z^2) - (1+b^2)xz\right)}$$

Theorem

Let A *be a nonsingular* $n \times n$ *matrix and* $\boldsymbol{c}, \boldsymbol{d} \in \mathbb{R}^n$. *Then*

1. $\det(\mathsf{I} + \boldsymbol{c}\boldsymbol{d}^T) = 1 + \boldsymbol{d}^T\boldsymbol{c}$,
2. $\det(\mathsf{A} + \boldsymbol{c}\boldsymbol{d}^T) = \det(\mathsf{A})(1 + \boldsymbol{d}^T\mathsf{A}^{-1}\boldsymbol{c})$.

Proof

1. The following identity "magically" provides the proof:

$$\begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \boldsymbol{d}^T & 1 \end{pmatrix} \begin{pmatrix} \mathsf{I} + \boldsymbol{c}\boldsymbol{d}^T & \boldsymbol{c} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ -\boldsymbol{d}^T & 1 \end{pmatrix} = \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \boldsymbol{d}^T & 1 \end{pmatrix} \begin{pmatrix} \mathsf{I} & \boldsymbol{c} \\ -\boldsymbol{d}^T & 1 \end{pmatrix} = \begin{pmatrix} \mathsf{I} & \boldsymbol{c} \\ \mathbf{0}^T & \boldsymbol{d}^T\boldsymbol{c} + 1 \end{pmatrix}$$

since

Proof

1. The following identity "magically" provides the proof:

$$\begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \boldsymbol{d}^T & 1 \end{pmatrix} \begin{pmatrix} \mathsf{I} + \boldsymbol{c}\boldsymbol{d}^T & \boldsymbol{c} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ -\boldsymbol{d}^T & 1 \end{pmatrix} = \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \boldsymbol{d}^T & 1 \end{pmatrix} \begin{pmatrix} \mathsf{I} & \boldsymbol{c} \\ -\boldsymbol{d}^T & 1 \end{pmatrix} = \begin{pmatrix} \mathsf{I} & \boldsymbol{c} \\ \mathbf{0}^T & \boldsymbol{d}^T\boldsymbol{c} + 1 \end{pmatrix}$$

since

$$\det \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \boldsymbol{d}^T & 1 \end{pmatrix} = 1, \qquad \det \begin{pmatrix} \mathsf{I} + \boldsymbol{c}\boldsymbol{d}^T & \boldsymbol{c} \\ \mathbf{0}^T & 1 \end{pmatrix} = \det(\mathsf{I} + \boldsymbol{c}\boldsymbol{d}^T),$$

$$\det \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ -\boldsymbol{d}^T & 1 \end{pmatrix} = 1, \qquad \det \begin{pmatrix} \mathsf{I} & \boldsymbol{c} \\ \mathbf{0}^T & \boldsymbol{d}^T\boldsymbol{c} + 1 \end{pmatrix} = 1 + \boldsymbol{d}^T\boldsymbol{c}.$$

Proof (cont.)

2. Rewrite

$$\mathsf{A} + \boldsymbol{c}\boldsymbol{d}^T = \mathsf{A}(\mathsf{I} + \mathsf{A}^{-1}\boldsymbol{c}\boldsymbol{d}^T),$$

Proof (cont.)

2. Rewrite

$$A + \boldsymbol{c}\boldsymbol{d}^T = A(I + A^{-1}\boldsymbol{c}\boldsymbol{d}^T),$$

let $\tilde{\boldsymbol{c}} = A^{-1}\boldsymbol{c}$ and then use (1) with $\tilde{\boldsymbol{c}}$ instead of $\boldsymbol{c}$, i.e.,

Proof (cont.)

2. Rewrite

$$A + \boldsymbol{c}\boldsymbol{d}^T = A(I + A^{-1}\boldsymbol{c}\boldsymbol{d}^T),$$

let $\tilde{\boldsymbol{c}} = A^{-1}\boldsymbol{c}$ and then use (1) with $\tilde{\boldsymbol{c}}$ instead of $\boldsymbol{c}$, i.e.,

$$\det(A + \boldsymbol{c}\boldsymbol{d}^T) = \det(A)\det(I + \underbrace{A^{-1}\boldsymbol{c}}_{=\tilde{\boldsymbol{c}}}\boldsymbol{d}^T)$$

Proof (cont.)

2. Rewrite
$$A + \boldsymbol{c}\boldsymbol{d}^T = A(I + A^{-1}\boldsymbol{c}\boldsymbol{d}^T),$$
let $\tilde{\boldsymbol{c}} = A^{-1}\boldsymbol{c}$ and then use (1) with $\tilde{\boldsymbol{c}}$ instead of $\boldsymbol{c}$, i.e.,
$$\det(A + \boldsymbol{c}\boldsymbol{d}^T) = \det(A)\det(I + \underbrace{A^{-1}\boldsymbol{c}}_{=\tilde{\boldsymbol{c}}}\boldsymbol{d}^T)$$
$$\overset{(1)}{=} \det(A)(1 + \boldsymbol{d}^T\tilde{\boldsymbol{c}})$$

Proof (cont.)

2. Rewrite
$$A + \boldsymbol{c}\boldsymbol{d}^T = A(I + A^{-1}\boldsymbol{c}\boldsymbol{d}^T),$$
let $\tilde{\boldsymbol{c}} = A^{-1}\boldsymbol{c}$ and then use (1) with $\tilde{\boldsymbol{c}}$ instead of $\boldsymbol{c}$, i.e.,
$$\begin{aligned}\det(A + \boldsymbol{c}\boldsymbol{d}^T) &= \det(A)\det(I + \underbrace{A^{-1}\boldsymbol{c}}_{=\tilde{\boldsymbol{c}}}\boldsymbol{d}^T) \\ &\overset{(1)}{=} \det(A)(1 + \boldsymbol{d}^T\tilde{\boldsymbol{c}}) \\ &= \det(A)(1 + \boldsymbol{d}^T A^{-1}\boldsymbol{c}).\end{aligned}$$

□

References I

[FM15] G. E. Fasshauer and M. J. McCourt, *Kernel-based Approximation Methods using* MATLAB, Interdisciplinary Mathematical Sciences, vol. 19, World Scientific Publishing, Singapore, 2015.

[FWL04] B. Fornberg, G. Wright, and E. Larsson, *Some observations regarding interpolants in the limit of flat radial basis functions*, Comput. Math. Appl. **47** (2004), 37–55, to appear.

[Sch05] Robert Schaback, *Multivariate interpolation by polynomials and radial basis functions*, Constr. Approx. **21** (2005), 293–317.