TMA 4115 Matematikk 3

MTEL, MTENERG, MTIØT, MTTK, II

Alexander Schmeding

NTNU

11. April 2014

The Gram-Schmidt Process

Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ be a basis for a non-zero subspace $W \subseteq \mathbb{R}^n$. Define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
\vdots &= \quad \vdots \qquad\qquad \vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \sum_{i=1}^{p-1} \frac{\mathbf{x}_p \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}\mathbf{v}_i
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W and in addition

$$\text{span}\ \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \text{span}\ \{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \quad \text{for } 1 \leq k \leq p$$

Application for Gram-Schmidt

(Spring 2011 Problem 5a)

Find an orthogonal basis of $\mathrm{Col}(A)$ with

$$A = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & 1 & 5 & -3 \\ -1 & -1 & -2 & 1 \end{bmatrix}.$$

To run Gram-Schmidt we need a basis of $\mathrm{Col}(A)$
($\rightarrow$ Gauss elimination on A!)

Alternatively, we can apply Gram-Schmidt to a generating system of $\mathrm{Col}(A)$. This is less work!
Let $\mathbf{x}_i$ be the i-th column of A.

Gram-Schmidt on a generating system

We note that all vectors in the generating set are non-zero. Thus set $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ then $\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$,

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \mathbf{0}$$

we get $\mathbf{0}$, so $\mathbf{x}_3$ was already contained in the span of $\mathbf{v}_1, \mathbf{v}_2$
$\rightsquigarrow$ discard $\mathbf{v}_3$ and continue with the next vector $\mathbf{x}_4$.

$$\mathbf{v}_3 = \mathbf{x}_4 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \mathbf{0}$$

again the result ist $\mathbf{0}$ and we discard $\mathbf{x}_4$.Since it was the last, the orthogonal basis of $\mathrm{Col}(A)$ is $\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$

An observation

In setting up the normal equation we computed matrices of the form $A^T A$. Here are some examples:

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}, \quad \begin{bmatrix} 4 & 17 \\ 17 & 81 \end{bmatrix}, \quad \begin{bmatrix} 6 & 6 & 12 & -6 \\ 6 & 11 & 7 & -1 \\ 12 & 7 & 29 & -17 \\ -6 & -1 & -17 & 11 \end{bmatrix}$$

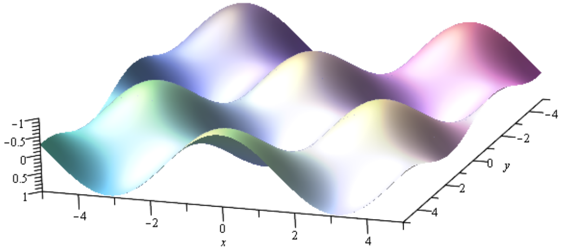
These matrices have an interesting structure:
Taking their transpose we get back the same matrix![1]

We call these special matrices **symmetric** and will now investigate their properties.

[1]Note that the rules for the transpose imply $(A^T A)^T = A^T (A^T)^T = A^T A$.

Application for quadratic forms

What are minima and maxima of
$f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto \cos(x)\sin(y)$?



Recall from calculus that for a (sufficiently) differentiable function $f\colon \mathbb{R} \to \mathbb{R}$ we know: If $f'(x) = 0$ and $f''(x) \neq 0$ then

if $f''(x) < 0$ we have a local maximum at x

if $f''(x) > 0$ we have a local minimum at x

For $f\colon \mathbb{R}^n \to \mathbb{R}$ (sufficiently) differentiable function a similar criterion holds with symmetric matrices!

Application for quadratic forms II

In calculus you learn how to compute for $\mathbf{x} \in \mathbb{R}^n$ the derivatives $f'(\mathbf{x})$ and $f''(\mathbf{x})$.

Furthermore, $f''(\mathbf{x})$ is determined by the **Hessian**

$$Hf(\mathbf{x}) = \left[\frac{\partial^2 f}{\partial_{x_i}\partial_{x_j}}(\mathbf{x})\right]_{1\leq i,j\leq n}$$

In the example $f\colon \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto \cos(x)\sin(y)$. The Hessian computes as:

$$Hf\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = -\begin{bmatrix} \cos(x)\sin(y) & \sin(x)\cos(y) \\ \sin(x)\cos(y) & \cos(x)\sin(y) \end{bmatrix}$$

It is a symmetric matrix[2].

Study maxima and minima via the Hessian!

[2] This is no coincidence, by Schwartz law the Hessian of each sufficiently differentiable function will be symmetric!

An important property

Definition A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called

- **positive definite**, if $Q(\mathbf{x}) > 0, \ \forall \mathbf{x} \neq \mathbf{0}$,
- **negative definite**, if $Q(\mathbf{x}) < 0, \ \forall \mathbf{x} \neq \mathbf{0}$,
- **indefinite**, if $Q(\mathbf{x})$ assumes both negative and positive values.

Theorem Let A be a symmetrix matrix and $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be the associated quadratic form.Then Q is

- positive definite, if all eigenvalues of A are positive
- negative definite, if all eigenvalues of A are negative
- indefinite, if there are positive and negative eigenvalues

Generalized criterion for minima and maxima

Let $f: \mathbb{R}^n \to \mathbb{R}$ be a sufficiently differentiable function.

Generalized criterion for minima and maxima:
If $f'(\mathbf{x}) = \mathbf{0}$ and the Hessian $Hf(\mathbf{x})$ is

- positive definite, then f has a local minimum at $\mathbf{x}$,
- negative definite, then f has a local maximum at $\mathbf{x}$,
- indefinite, then f has a saddle point at $\mathbf{x}$.