Math primer for Control theory TSRT09

Daniel Axehill, Isak Nielsen, and Sina Khoshfetrat Pakazad

January 23, 2014

This document includes some basic mathematical concepts that will be used during the course TSRT09. The notation is inherited from the course book "Reglerteori" by Glad & Ljung. For a more thorough description, the reader is referred to standard linear algebra literature and the course material.

Scalar products

The standard *inner product* (or *dot product* or *scalar product*) between two vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ is defined as

$$x \cdot y = x^T y = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i. \tag{1}$$

Two vectors are *perpendicular* to each other if and only if their inner product is zero, i.e.,

$$x \perp y \iff x^T y = 0, \tag{2}$$

where $x \perp y$ means that x and y are orthogonal.

Matrix-vector products

A *matrix-vector product* between a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $x \in \mathbb{R}^n$ results in a new vector $z \in \mathbb{R}^m$, where the i-th element (row) in z is obtained as the inner product between row i in A and the vector x. For $m = 2$ and $n = 3$ the following result is obtained

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{bmatrix}. \tag{3}$$

Norms and Cauchy-Schwartz inequality

A *norm* on a finite dimensional vector space (e.g. $\mathbb{R}^n$) is denoted $|x|$, where $|\cdot|$ is any function satisfying the following properties for all x in the vector space:

1. $|x| \geq 0$
2. $|x| = 0 \iff x = 0$
3. $|\alpha x| = |\alpha|\,|x|$
4. $|x + y| \leq |x| + |y|$ (the so called *triangle inequality*)

For the vector space $\mathbb{R}^n$, commonly used norms are the 2-norm ($|x|_2$), the 1-norm ($|x|_1$) and the ∞-norm ($|x|_\infty$).

For an inner product space such as $\mathbb{R}^n$, the inner product and the norm are related by

$$|x|_2 = \sqrt{x^T x},$$

where $x^T x$ is the inner product of x with itself. Furthermore, for all $x, y \in \mathbb{R}^n$ the so called *Cauchy-Schwartz inequality* holds

$$|x^T y| \leq |x|\,|y|. \tag{4}$$

A norm on a infinite dimensional functional space (e.g. $\mathcal{H}_2$) is denoted $||x||$. It is defined in the same way as $|x|$, and satisfies the same properties as $|x|$, and the double bars are just for notational convenience.

Determinants

The *determinant*, denoted $\det(A)$, is a scalar value associated with a square matrix $A \in \mathbb{R}^{n \times n}$. For a 2×2 matrix, the determinant is computed as

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \tag{5}$$

and for a 3×3 matrix it can be computed as e.g.

$$\begin{aligned} \det(A) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned} \tag{6}$$

If $\det(A) = 0$ the matrix A is *singular* or *non-invertible*.

Eigenvalues and eigenvectors

Consider a square matrix A. Then λ is called an *eigenvalue* and $v \neq 0$ an *eigenvector* to A if they satisfy

$$Av = \lambda v \tag{7}$$

holds. Hence, the result of the matrix-vector product Av is a scaled version of v, and the scaling is called the eigenvalue. The eigenvalues can be calculated using the *characteristic equation* (or *secular equation*)

$$\det(A - \lambda I) = 0 \tag{8}$$

The corresponding eigenvectors can be found by inserting the calculated eigenvalues one by one in (7), i.e.,

$$(A - I\lambda_i)v_i = 0 \tag{9}$$

where λ_i is eigenvalue i that has been computed from the characteristic equation above and v_i is the corresponding (non-zero) eigenvector.

Two sometimes useful identities involving eigenvalues are

$$\operatorname{trace}(A) = \sum_{i=1}^{n} \lambda_i \tag{10}$$

$$\det(A) = \prod_{i=1}^{n} \lambda_i \tag{11}$$

As a result, a singular matrix must have at least one zero eigenvalue.

Null space

The *nullspace* (or *kernel*), denoted $\mathcal{N}(A)$, of a matrix $A \in \mathbb{R}^{m \times n}$ consists of all vectors $x \in \mathbb{R}^n$ such that $Ax = 0$, i.e,

$$\mathcal{N}(A) = \{x \,|\, Ax = 0,\ x \in \mathbb{R}^n\}. \tag{12}$$

Hence, the nullspace consists of all vectors $x \in \mathbb{R}^n$ that are perpendicular to each row of A. This is clear by looking at

$$Ax = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ \vdots \\ a_m^T x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \iff a_i^T x = 0,\ i = 1, \dots, m, \tag{13}$$

where a_i^T, $i = 1, \dots, m$ are the rows of A, and $a_i^T x = 0$ means that $a_i \perp x$.

Range space

The *range space* (or *column space*), denoted $\mathcal{R}(A)$, of a matrix $A \in \mathbb{R}^{m\times n}$ consists of all mappings of A, i.e.,

$$\mathcal{R}(A) = \{y \in \mathbb{R}^m \,|\, y = Ax,\ \forall x \in \mathbb{R}^n\}. \tag{14}$$

Hence the range space of A consists of all vectors that are linear combinations of the columns of A. This is clear by looking at

$$y = Ax = \begin{bmatrix}\bar{a}_1 & \dots & \bar{a}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix} = \sum_{i=1}^{n} x_i\bar{a}_i, \tag{15}$$

where $\bar{a}_i,\ i = 1,\dots,n$ are the columns of A. Suppose that there are r linearly independent columns in A, then these are a basis for the range space and the dimension of the range space of A is $n_{\mathcal{R}}$.

Rank

The number of linearly independent columns of a matrix A is called the *column rank*, whereas the number of linearly independent rows are called the *row rank*. The column rank is equal to the row rank, and this value is referred to as the *rank* of the matrix A.

The column rank (and hence the rank) of A is the same as the dimension of the range space of A, i.e.,

$$\text{rank}(A) = \dim(\mathcal{R}(A)) = r, \tag{16}$$

where r is the number of linearly independent columns or rows in the matrix A. A square matrix (i.e. $A \in \mathbb{R}^{n\times n}$) is invertible (or non-singular) if and only if the matrix has full rank (i.e. $\text{rank}(A) = n$).

Positive (semi) definite matrices and quadratic forms

Let $\mathbb{S}^n$ denote a real symmetric matrix with n rows (and columns). For a symmetric matrix it holds that the quadratic form (x^TAx) can be bounded from below and from above as

$$\lambda_{min}(A)x^Tx \le x^TAx \le \lambda_{max}(A)x^Tx \tag{17}$$

where $\lambda_{min}(A)$ denotes the smallest eigenvalue of A and $\lambda_{max}(A)$ denotes the largest one. A symmetric matrix $A \in \mathbb{S}^n$ is called *positive definite* if for all $x \neq 0$ it holds that $x^TAx > 0$. This is denoted as $A \succ 0$ and this is true if and only if all eigenvalues of A are strictly positive. The set of positive definite matrices is commonly denoted as $\mathbb{S}^n_{++}$.

Furthermore, a matrix $A \in \mathbb{S}^n$ for which is holds that $x^TAx \ge 0$ for all $x \neq 0$ is called positive semidefinite. This is denoted as $A \succeq 0$ and this is true if and only if all eigenvalues of A are non-negative (positive or zero). The set of positive semidefinite matrices is commonly denoted as $\mathbb{S}^n_{+}$.

Singular value decomposition

For every matrix $A \in \mathbb{R}^{m\times n}$ with $\text{rank}(A) = r \le \min(m,n)$, A can be factored as

$$A = \underbrace{\begin{bmatrix}\bar{U} & \tilde{U}\end{bmatrix}}_{U}\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}\underbrace{\begin{bmatrix}\bar{V} & \tilde{V}\end{bmatrix}^T}_{V^T}, \tag{18}$$

where $\Sigma = \text{diag}(\sigma_1,\dots,\sigma_r)$ with

$$\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_r > 0$$

and $\bar{U} \in \mathbb{R}^{m\times r}$, $\tilde{U} \in \mathbb{R}^{m\times m-r}$, $\bar{V} \in \mathbb{R}^{n\times r}$, $\tilde{V} \in \mathbb{R}^{n\times n-r}$ with $U^TU = I_m$ and $V^TV = I_n$. The matrices I_m and I_n denote $m \times m$ and $n \times n$ identity matrices, respectively. The factorization in (18) is called the *singular value decomposition* of A and the numbers σ_i are the *singular values*. Note that columns of $\bar{U}$ are a basis for the range space of A and columns of $\tilde{V}$ are a basis of null space of A. Here we assumed that A is real-valued, however, the same definitions hold even if A is complex. In this case, matrices U and V can be complex-valued and wherever we use matrix transpose, i.e., ^{T}, it should be replaced by matrix conjugate-transpose, i.e., *. However, the singular values σ_i are always real valued.

Matrix inverse

A matrix $A \in \mathbb{R}^{n\times n}$ is *invertible* or *non-singular* if there exists a matrix $A^{-1} \in \mathbb{R}^{n\times n}$ such that

$$A^{-1}A = AA^{-1} = I. \tag{19}$$

For a 2×2 matrix A the inverse is computed as

$$\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}^{-1} = \frac{1}{\det(A)}\begin{bmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{bmatrix} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}}\begin{bmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{bmatrix} \tag{20}$$

A *singular* matrix $B \in \mathbb{R}^{n\times n}$ is a matrix that is not invertible. A singular matrix B always have $\det(B) = 0$, and at least one eigenvalue is zero and hence there exist a null space to B.

For singular matrices and non-square matrices, the *(Moore-Penrose) pseudo inverse*, denoted by †, is a generalization of the matrix inverse. The pseudo inverse $A^\dagger$ of the matrix $A \in \mathbb{R}^{m\times n}$ is unique and satisfies the four properties

$$AA^\dagger A = A \tag{21}$$

$$A^\dagger AA^\dagger = A^\dagger \tag{22}$$

$$(AA^\dagger)^* = AA^\dagger \tag{23}$$

$$(A^\dagger A)^* = A^\dagger A. \tag{24}$$

If A is invertible, then $A^\dagger = A^{-1}$.

The pseudo inverse can be computed in different ways, e.g. using singular value decompositions. In MATLAB it can be computed using the command `pinv`.

Jacobians and chain rule

Consider an m-dimensional vector-valued function $f : \mathbb{R}^n \to \mathbb{R}^m$. The Jacobian matrix for this function is defined as

$$J_f(x) = \begin{bmatrix}\frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \dots & \frac{\partial f_1(x)}{\partial x_n}\\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \dots & \frac{\partial f_n(x)}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \dots & \frac{\partial f_n(x)}{\partial x_n}\end{bmatrix}, \tag{25}$$

which can be used to define the first order approximation of the function around x that is given as

$$f(z) \approx f(x) + J_f(x)(z - x) \tag{26}$$

for every point z that is close enough to x. Now consider a vector-valued function $h : \mathbb{R}^n \to \mathbb{R}^p$ that is defined as $h(x) = g(f(x))$ where $g : \mathbb{R}^m \to \mathbb{R}^p$ is another vector-valued function. The Jacobian matrix for this function can be computed using the chain rule which states that

$$J_h(x) = J_g(f(x))J_f(x). \tag{27}$$

Note: It is *not* allowed to bring this document to an exam in TSRT09!