Reminders about matrix manipulations

In class I went maybe too quickly over some matrix operations. I would like here to detail which manipulations are legit. **Note that this document is by no mean exhaustive**.

Matrix multiplication

When I write that A is a $m \times n$ matrix, it means that it has m row and n columns. My vectors are always columns vectors, that is $\mathbf{c} \in \mathbb{R}^n$ means that $\mathbf{c}$ is seen as $n \times 1$ matrix. **So in the formulas below, a vector is just a particular case of matrix with only one column**. The transpose of the matrix A of size $m \times n$, denoted by $A^\top$ is a $n \times m$ matrix: if $A = (A_{ij})_{ij}$ then $A^\top = (A_{ji})_{ij}$. In particular, **the transpose of a column vector is a row vector**. Also, the transposition is an involution in the sense that $(A^\top)^\top = A$.

If A is a $m_1 \times n_1$ matrix and B is a $m_2 \times n_2$ matrix, you can define the product AB if and only if $n_1 = m_2$ (the "interior" dimensions coincide) and the result is a $m_1 \times n_2$ matrix (the "exterior" dimensions).

The definition of matrix multiplications are recalled in a document that you can find on Anstee's website[1]. As an example, if

$$A = \begin{pmatrix} 2 & 5 \\ -3 & 8 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 0 & 1 & -1 \\ 0 & 2 & 1 & -1 \end{pmatrix}, \quad \text{then} \quad AB = \begin{pmatrix} 10 & 10 & 7 & -7 \\ -15 & 16 & 5 & -5 \\ 0 & 2 & 1 & -1 \end{pmatrix}.$$

As a particular case for dimensions, if A is a $m \times n$ matrix, $\mathbf{c}$ is a $n \times 1$ vector and $\mathbf{b}$ is a $m \times 1$ vector (both columns vector), then $A\mathbf{c}$ is a $m \times 1$ column vector whereas $\mathbf{b}^\top A$ is a $1 \times n$ row vector. In other words, **multiplying a matrix on the right by a column vector gives you a column vector, multiplying a matrix on the left by a row vector gives you a row vector**.

Below are some rules about matrix multiplication which are always valid, provided that the dimensions of the matrix involved are compatible.

- *Associativity*. There holds $A(BC) = (AB)C$ which justifies the notations ABC (without parenthesis). Don't pay too much attention to this rule, you use it without even realizing it most of the time.
- *Non commutativity*. In general, $AB \neq BA$. That is the order in a product matter. Actually, if the dimensions are such that AB makes sense, in general BA doesn't make sense.
- *Bilinearity*. This is a technical word to say that you can expand and factor matrix products as long as you pay attention to non commutativity. That is,

  $$(A + B)(C + D) = AC + AD + BC + BD.$$

  Be careful that for instance $(A + B)^2 = A^2 + AB + BA + B^2$ but we can't go no further because $AB \neq BA$ in general. Actually, if $\alpha, \beta, \gamma, \delta$ are scalars (real numbers) then

  $$(\alpha A + \beta B)(\gamma C + \delta D) = \alpha\gamma AC + \alpha\delta AD + \beta\gamma BC + \beta\delta BD,$$

  where we recall that multiplying a matrix by a scalar multiplies *all* its entries by the same scalar.

[1] `https://www.math.ubc.ca/~anstee/math340/340matrixmult.pdf`

- *Transposition and product*. A rule that I have used intensively is that
$$(AB)^\top = B^\top A^\top,$$
that is transposition reverses the order in a product. This rule is the one which makes sense in term of dimensions: if A is $m \times n$ and B is $n \times p$, then $(AB)^\top$ is $p \times m$ and I let you check that this is the same as the dimension of $B^\top A^\top$. This rule can be extended to a product of more matrices:
$$(A_1 A_2 \dots A_{k-1} A_k)^\top = A_k^\top A_{k-1}^\top \dots A_2^\top A_1^\top.$$
- *Transposition and inverse*. A corollary of the previous rule is that if A is a square matrix which is invertible then
$$\left(A^\top\right)^{-1} = \left(A^{-1}\right)^\top$$
and we call this matrix $A^{-\top}$: transposition and inverse commute.
- *Inverse and product*. A rule that I haven't used much but which is good to know: if A and B are square matrices of the same dimension and are invertible,
$$(AB)^{-1} = B^{-1} A^{-1}.$$
In other words, taking the inverse change the order of the product. This rule can be extended to a product of more matrices:
$$(A_1 A_2 \dots A_{k-1} A_k)^{-1} = A_k^{-1} A_{k-1}^{-1} \dots A_2^{-1} A_1^{-1}.$$

Block matrices

In the proof of the revised simplex formulas I used block matrices. **The only thing that you need to know is that the same rules than for mutliplication apply, you just have to keep track of non commutativity** (that is the order of the multiplications matter). That is, imagine that you have a $m \times n$ matrix A and a $n \times p$ matrix B that you can write
$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \qquad \text{then} \qquad B = \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right)$$
with compatible dimensions: A_{11} is $m_1 \times n_1$, A_{12} is $m_1 \times n_2$, A_{21} is $m_2 \times n_1$ and A_{22} is $m_2 \times n_2$ (with $n_1 + n_2 = n$ and $m_1 + m_2 = m$) while B_{11} is $n_1 \times p_1$, B_{12} is $n_1 \times p_2$, B_{21} is $n_2 \times p_1$ and B_{22} is $n_2 \times p_2$ (with $p_1 + p_2 = p$). (This looks scary but it just means that the number of columns of A_{11} is the same as the number of line of B_{11} and B_{12} for instance). Then you can apply the "usual" product rule and write
$$AB = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right) = \left(\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array}\right).$$
The only thing you should take care of is that $A_{11}B_{11} \neq B_{11}A_{11}$ (this is non commutativity) and so on for all products between blocks of A and B.

Same for transposition! For instance,
$$\text{if} \quad A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \qquad \text{then} \quad A^\top = \left(\begin{array}{c|c} A_{11}^\top & A_{21}^\top \\ \hline A_{12}^\top & A_{22}^\top \end{array}\right).$$

Scalar product

The (canonical) scalar product of two vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$ of the same dimension is defined as
$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{y}_i.$$

For instance, if

$$\mathbf{x} = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} -1 \\ -5 \\ 1 \end{pmatrix}, \quad \text{then} \quad \mathbf{x} \cdot \mathbf{y} = 3 \times (-1) + (-2) \times (-5) + 0 \times 1 = 7.$$

The scalar product is a real number.

Now, because of how we define matrix multiplication and transposition, there holds

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\top \mathbf{y}$$

provided $\mathbf{x}$ and $\mathbf{y}$ are recorded as $n \times 1$ vectors.

Indeed, in this case $\mathbf{x}^\top$ is a $1 \times n$ vector so $\mathbf{x}^\top \mathbf{y}$ is a 1×1 matrix, that is a scalar.

I have often used that a 1×1 matrix is its own transpose.. For instance in that case, as $\mathbf{x}^\top \mathbf{y}$ is a 1×1 matrix

$$\mathbf{x}^\top \mathbf{y} = (\mathbf{x}^\top \mathbf{y})^\top = \mathbf{y}^\top (\mathbf{x}^\top)^\top = \mathbf{y}^\top \mathbf{x}$$

because transposing twice gives the original matrix. The right hand side is $\mathbf{y} \cdot \mathbf{x}$ by definition, hence I have "proved" that the scalar product is a symmetric function of $\mathbf{x}$ and $\mathbf{y}$. **In the proofs, I ended up transposing a 1×1 matrix when I wanted to use the symmetry of scalar product because I decided to use only the transpose notation and not the scalar product one**.

An expression which was quite recurrent was the following. Let A be a $m \times n$ matrix, $\mathbf{x}$ be a $n \times 1$ vector and $\mathbf{y}$ be a $m \times 1$ vector. Then I let you check the following chains of equalities:

$$\mathbf{y} \cdot (A\mathbf{x}) = \mathbf{y}^\top A\mathbf{x} = (\mathbf{y}^\top A\mathbf{x})^\top = \mathbf{x}^\top A^\top \mathbf{y} = \mathbf{x} \cdot (A^\top \mathbf{y}),$$

and these are equalities between 1×1 matrices (a.k.a. real numbers). As an example $n = 3, m = 2$, and

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\text{then} \quad \mathbf{y}^\top A\mathbf{x} = x_1 y_1 a + x_2 y_1 b + x_3 y_1 c + x_1 y_2 d + x_2 y_2 e + x_3 y_2 f$$

Do the matrix product yourself to understand the structure of the right hand side!

Inequalities between vectors

Last aspect, about inequalities. I recall that we have defined, for two vectors of the same dimension n, that $\mathbf{x} \leqslant \mathbf{y}$ if for every index i there holds $\mathbf{x}_i \leqslant \mathbf{y}_i$. For instance, if

$$\mathbf{x} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 5 \\ -1 \\ 19 \end{pmatrix} \quad \text{then} \quad \mathbf{x} \leqslant \mathbf{y}$$

but if

$$\mathbf{x} = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 4 \\ -1 \\ 7 \end{pmatrix} \quad \text{then there doesn't hold} \quad \mathbf{x} \leqslant \mathbf{y}$$

because of the second component ($0 > -1$).

Now, what can you do with inequalities between vectors?

- *Transpose them.* If $\mathbf{x} \leqslant \mathbf{y}$ then $\mathbf{x}^\top \leqslant \mathbf{y}^\top$, that is *transposition doesn't change the inequality.* (Actually one could say that this is almost by definition).
- *Multiply by a scalar.* Let's assume that $\mathbf{x} \leqslant \mathbf{y}$. Then, for $\alpha > 0$ there holds $\alpha\mathbf{x} \leqslant \alpha\mathbf{y}$ while for $\alpha < 0$ there holds $\alpha\mathbf{x} \geqslant \alpha\mathbf{y}$. This is the same as for usual inequalities, multiplying by positive numbers keep the inequality unchanged, for a negative number you have to change the sign.
- *Take the scalar product with a positive vector.* It means that

$$\text{if} \quad \mathbf{x} \leqslant \mathbf{y} \quad \text{and} \quad \mathbf{z} \geqslant \mathbf{0} \quad \text{then} \quad \mathbf{z}^\top\mathbf{x} \leqslant \mathbf{z}^\top\mathbf{y}.$$

provided that $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are columns vectors of the same dimension. Notice that $\mathbf{z}^\top\mathbf{x}$ and $\mathbf{z}^\top\mathbf{y}$ are scalar products, hence they are scalars: the last inequality is an inequality between real numbers. **In short, you can take the scalar product of a vectorial inequality with a non-negative vector**.

To understand why the last property is true, let's check it with an example. Take

$$\mathbf{x} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 5 \\ -1 \\ 19 \end{pmatrix} \quad \text{and} \quad \mathbf{z} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

We have $\mathbf{z} \geqslant \mathbf{0}$ and also $\mathbf{x} \leqslant \mathbf{y}$: we write the latter as the system of inequalities

$$\begin{cases} 2 & \leqslant & 5 \\ -3 & \leqslant & -1 \\ 1 & \leqslant & 19 \end{cases}$$

Then I multiply the first inequality by $\mathbf{z}_1 = 1$, the second one by $\mathbf{z}_2 = 2$ and the third one by $\mathbf{z}_3 = 3$, and I sum everything. As all the coefficients of $\mathbf{z}$ are non negative I do not change the signs of inequalities hence

$$-1 = 1 \times 2 + 2 \times (-3) + 3 \times 1 \leqslant 1 \times 5 + 2 \times (-1) + 3 \times 19 = 60.$$

I leave you to check that the left hand side is $\mathbf{z}^\top\mathbf{x}$ while the right hand side is $\mathbf{z}^\top\mathbf{y}$.

Example: proof of weak duality Let me take advantage of the notations above to give a cleaner proof of weak duality. For that I recall that A is a $m \times n$ matrix, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^n$. The statement of weak duality looks like:

If $\mathbf{x}$ and $\mathbf{y}$ satisfy

$$\begin{cases} A\mathbf{x} & \leqslant & \mathbf{b} \\ \mathbf{x} & \geqslant & \mathbf{0} \end{cases} \quad \text{and} \quad \begin{cases} A^\top\mathbf{y} & \geqslant & \mathbf{c} \\ \mathbf{y} & \geqslant & \mathbf{0} \end{cases}$$

then there holds

$$\mathbf{c}^\top\mathbf{x} \leqslant \mathbf{b}^\top\mathbf{y}.$$

Now to prove it I just write the following chain of inequalities:

$$\mathbf{c}^\top\mathbf{x} = \mathbf{x}^\top\mathbf{c} \leqslant \mathbf{x}^\top(A^\top\mathbf{y}) = \mathbf{y}^\top A\mathbf{x} \leqslant \mathbf{y}^\top\mathbf{b} = \mathbf{b}^\top\mathbf{y}.$$

Indeed, all the equalities come from symmetry of the scalar product, while the first inequality is obtained from taking the scalar product of $A^\top\mathbf{y} \geqslant \mathbf{c}$ with $\mathbf{x} \geqslant \mathbf{0}$, and the second inequality from taking the scalar product of $A\mathbf{x} \leqslant \mathbf{b}$ with $\mathbf{y} \geqslant \mathbf{0}$.

Exercises

Disclaimer: I don't claim that these are the most relevant exercises to prepare for the midterm or final, just that it's related to the content of these notes. I think they are harder than what you can be asked in midterm of final. I haven't typed up solutions, but please feel free to send me an email or come to an office hour if you have questions.

1. Go back to the proof of the revised simplex formulas and the strong duality theorem and try to understand them.

2. Prove that for any invertible matrix A of size $n \times n$,
$$\left(A^{-1}\right)^{\top} = \left(A^{\top}\right)^{-1}.$$
(*Hint*. To prove that B is the inverse of C, it is enough to show $BC = \mathrm{Id}_n$, provided B and C are of size $n \times n$).

3. Let A a $n \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$ a column vector. Show that
$$\mathbf{x}^{\top} A^{\top} A \mathbf{x} \geqslant 0.$$

4. Let A, B, C, D matrices such that the dimensions in the equations below match. Let $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v}$ columns vectors as well as $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ other columns vectors such that
$$\left\{\begin{array}{llll} A\mathbf{x} & +B\mathbf{y} & = & \mathbf{c} \\ C\mathbf{x} & +D\mathbf{y} & \leqslant & \mathbf{d} \\ \mathbf{x} & & \geqslant & \mathbf{0} \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{llll} A^{\top}\mathbf{u} & +C^{\top}\mathbf{v} & \geqslant & \mathbf{a} \\ B^{\top}\mathbf{u} & +D^{\top}\mathbf{v} & = & \mathbf{b} \\ & \mathbf{v} & \geqslant & \mathbf{0} \end{array}\right. .$$
Prove that
$$\mathbf{a}^{\top}\mathbf{x} + \mathbf{b}^{\top}\mathbf{y} \leqslant \mathbf{c}^{\top}\mathbf{u} + \mathbf{d}^{\top}\mathbf{v}$$
(*Comment*. This is a weak duality result for a LP not in standard form, see Question 2 of Assignment 1.)

5. Let A a matrix with the block decomposition
$$A = \left(\begin{array}{c|c} B & C \\ \hline \mathbf{0} & D \end{array}\right),$$
where B is of size $n \times n$, D of size $m \times m$ while C is of size $n \times m$ and $\mathbf{0}$ is the matrix full of zeros of size $m \times n$. In particular, A is of size $(n+m) \times (n+m)$. Prove that if B and D are invertible, so is A and compute A^{-1}.
(*Hint*. What happens if $n = m = 1$, that is A is a 2×2 matrix? Use this to guess an expression for A^{-1} and then check that it works with block matrices product.)