# Crimes Against Matrices

In this note we list some matrix crimes that we have, sadly, witnessed too often. Be very careful to avoid committing any of these crimes; in ENGR108 we have a *zero-tolerance* policy for *crimes against matrices*, at least on things you hand in to us. (What you do with matrices in your spare time, or on scratch paper, is of course your own business. But we recommend you avoid these crimes at all times, in order to not build bad habits.)

Check your work — don't become just another sad statistic!

## Syntax crimes

In a syntax crime, the perpetrator attempts to combine matrices (or other mathematical objects) in ways that violate basic syntax rules. These are serious crimes of negligence, since it is so easy to check your work for potential violations. We list some typical examples below.

- Adding, subtracting, or equating matrices (or vectors) of different dimensions.

  *Example:* writing $A + B$, when $A$ is a $2 \times 3$ matrix and $B$ is a $3 \times 3$ matrix.

- Violating the rules of constructing block matrices (*e.g.*, the submatrices in any row of a block matrix must have the same number of rows).

  *Example:* forming the block matrix $[A\ B]$, when $A$ is a $2 \times 3$ matrix and $B$ is a $3 \times 3$ matrix.

- Multiplying matrices with incompatible dimensions (*i.e.*, forming $AB$, when the number of columns of $A$ does not equal the number of rows of $B$).

  *Example:* forming $A^T B$, when when $A$ is a $2 \times 3$ matrix and $B$ is a $3 \times 3$ matrix.

- Taking the inverse or powers of a nonsquare matrix.

  *Example:* forming $A^{-1}$, when $A$ is a $2 \times 3$ matrix.

## Semantic crimes

In a semantic crime, the perpetrator forms an expression or makes an assertion that does not break any syntax rules, but is wrong because of the meaning. These crimes are a bit harder to detect than syntax crimes, so you need to be more vigilant to avoid committing them.

- Taking the inverse of a square, but non-invertible matrix. (Taking the inverse of a nonsquare matrix is a syntax crime—see above.)

  *Example:* forming $(ww^T)^{-1}$, where $w$ is a 2-vector.

  *Note:* writing $(ww^T)^{-1} = (w^T)^{-1}w^{-1}$, when $w$ is a 2-vector, involves both a syntax and semantic crime.

- Referring to a left inverse of a strictly wide matrix or a right inverse of a strictly tall matrix.

  *Example:* writing $QQ^T = I$, when $Q$ is a $5 \times 3$ matrix.

- Cancelling matrices on the left or right in inappropriate circumstances, *e.g.*, concluding that $B = C$ from $AB = AC$, when $A$ is not known to have independent columns.

  *Example:* concluding $x = y$ from $a^T x = a^T y$, when $a,\ x,\ y$ are 4-vectors.

## Miscellaneous crimes

Some crimes are hard to classify, or involve both syntax and semantic elements. Incorrect use of a matrix identity often falls in this class.

- Using $(AB)^T = A^T B^T$ (instead of the correct formula $(AB)^T = B^T A^T$).

  *Note:* this also violates syntax rules, if $A^T B^T$ is not a valid product.

- Using $(AB)^{-1} = A^{-1}B^{-1}$ (instead of the correct formula $(AB)^{-1} = B^{-1}A^{-1}$).

  *Note:* $(AB)^{-1} = A^{-1}B^{-1}$ violates syntax rules, if $A$ or $B$ is not square; it violates semantic rules if $A$ or $B$ is not invertible.

- Using $(A+B)^2 = A^2 + 2AB + B^2$. This (false) identity relies on the very useful, but unfortunately false, identity $AB = BA$.

## An example

Let's consider the expression $(A^T B)^{-1}$, where $A$ is a $m \times n$ matrix and $B$ is a $k \times p$ matrix. Here's how you might check for various crimes you might commit in forming this expression.

- We multiply $A^T$, which is $n \times m$, and $B$, which is $k \times p$, so we better have $m = k$ to avoid a syntax violation.

  *Note:* if $A$ is a scalar, then $A^T B$ might be a strange thing to write, but can be argued to not violate syntax, even though $m \neq k$. In a similar way, when $B$ is a scalar, you can write $A^T B$, and argue that syntax is not violated.

- The product $A^T B$ is $n \times p$, so we better have $n = p$ in order to (attempt to) invert it. At this point, we know that the dimensions of $A$ and $B$ must be the same (ignoring the case where one or the other is interpreted as a scalar).
- If $A^T B$ is a strictly tall–strictly wide product (*i.e.*, $A$ and $B$ are strictly wide), then $A^T B$ cannot possibly be invertible, so we have a semantic violation. To avoid this, we must have $A$ and $B$ square or tall, *i.e.*, $m \geq n$.

Summary: to write $(A^T B)^{-1}$ (assuming neither $A$ nor $B$ is interpreted as a scalar), $A$ and $B$ must have the same dimensions, and be tall or square.

Of course, even if $A$ and $B$ have the same dimensions, and are tall or square, the matrix $A^T B$ can be singular, in which case $(A^T B)^{-1}$ is meaningless. The point of our analysis above is that if $A$ and $B$ don't have the same dimension, or if $A$ and $B$ are strictly wide, then $(A^T B)^{-1}$ is *guaranteed* to be meaningless, no matter what values $A$ and $B$ might have in your application or argument.