ST552, Homework 1

Due Wednesday, Sep 4, 2013

Please make your proofs self-contained, without citing any corresponding theorem in the textbook or lecture notes

1. Show that for an arbitrary matrix $\boldsymbol{A}$, the maximum number of linearly independent rows equals the maximum number of linearly independent columns. Therefore the rank can be defined either way.

2. Assume $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n \times p}$. Show the following facts about the effect of matrix multiplication on the rank.

   (a) $\text{rank}(\boldsymbol{AB}) \leq \min\{\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B})\}$ for any $\boldsymbol{B}$.

   (b) $\text{rank}(\boldsymbol{AB}) = \text{rank}(\boldsymbol{B})$ for any $\boldsymbol{A}$ of full column rank.

   (c) $\text{rank}(\boldsymbol{AB}) = \text{rank}(\boldsymbol{A})$ for any $\boldsymbol{B}$ of full row rank.

   (d) $\text{rank}(\boldsymbol{AA}^T) = \text{rank}(\boldsymbol{A}^T\boldsymbol{A}) = \text{rank}(\boldsymbol{A})$. This is same as 3(c) below.

   (e) $\text{rank}(\boldsymbol{AA}^-) = \text{rank}(\boldsymbol{A}^-\boldsymbol{A}) = \text{rank}(\boldsymbol{A})$.

3. Show the following facts about the *Gramian* matrix $\boldsymbol{A}^\intercal\boldsymbol{A}$.

   (a) $\boldsymbol{A}^\intercal\boldsymbol{A}$ is symmetric and positive semidefinite.

   (b) $\mathcal{C}(\boldsymbol{A}) = \mathcal{C}(\boldsymbol{AA}^\intercal)$.

   (c) $\text{rank}(\boldsymbol{A}) = \text{rank}(\boldsymbol{A}^\intercal) = \text{rank}(\boldsymbol{A}^\intercal\boldsymbol{A}) = \text{rank}(\boldsymbol{AA}^\intercal)$.

   (d) $\boldsymbol{A}^\intercal\boldsymbol{A} = \boldsymbol{0}$ if and only if $\boldsymbol{A} = \boldsymbol{0}$.

   (e) $\boldsymbol{BA}^\intercal\boldsymbol{A} = \boldsymbol{CA}^\intercal\boldsymbol{A}$ if and only if $\boldsymbol{BA}^\intercal = \boldsymbol{CA}^\intercal$.

   (f) $\boldsymbol{A}^\intercal\boldsymbol{AB} = \boldsymbol{A}^\intercal\boldsymbol{AC}$ if and only if $\boldsymbol{AB} = \boldsymbol{AC}$.

   (g) For any generalized inverse $(\boldsymbol{A}^\intercal\boldsymbol{A})^-$, $[(\boldsymbol{A}^\intercal\boldsymbol{A})^-]^\intercal$ is also a generalized inverse of $\boldsymbol{A}^\intercal\boldsymbol{A}$. Note $(\boldsymbol{A}^\intercal\boldsymbol{A})^-$ is not necessarily symmetric.

   (h) $(\boldsymbol{A}^\intercal\boldsymbol{A})^-\boldsymbol{A}^\intercal$ is a generalized inverse of $\boldsymbol{A}$.

   (i) $\boldsymbol{AA}^+ = \boldsymbol{A}(\boldsymbol{A}^\intercal\boldsymbol{A})^-\boldsymbol{A}^\intercal$, where $\boldsymbol{A}^+$ is the Moore-Penrose inverse of $\boldsymbol{A}$.

   (j) Let $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{A}^\intercal\boldsymbol{A})^-\boldsymbol{A}^\intercal$. Show that $\boldsymbol{P}_{\boldsymbol{A}}$ is symmetric, idempotent, invariant to the choice of generalized inverse $(\boldsymbol{A}^\intercal\boldsymbol{A})^-$, and projects onto $\mathcal{C}(\boldsymbol{A})$.

4. (a) Show the Sherman-Morrison formula

$$(\boldsymbol{A} + \boldsymbol{uu}^\intercal)^{-1} = \boldsymbol{A}^{-1} - \frac{1}{1 + \boldsymbol{u}^\intercal\boldsymbol{A}^{-1}\boldsymbol{u}}\boldsymbol{A}^{-1}\boldsymbol{uu}^\intercal\boldsymbol{A}^{-1},$$

   where $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is a nonsingular matrix and $\boldsymbol{u} \in \mathbb{R}^n$. This formula supplies the inverse of the symmetric, rank-one perturbation of $\boldsymbol{A}$.

(b) Show the Woodbury formula

$$(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{V}^\intercal)^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}(\boldsymbol{I}_m + \boldsymbol{V}^\intercal\boldsymbol{A}^{-1}\boldsymbol{U})^{-1}\boldsymbol{V}^\intercal\boldsymbol{A}^{-1},$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is nonsingular, $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{R}^{n\times m}$, and $\boldsymbol{I}_m$ is the $m \times m$ identity matrix. In many applications m is much smaller than n. Woodbury formula generalizes Sherman-Morrison and is valuable because the smaller matrix $\boldsymbol{I}_m + \boldsymbol{V}^\intercal\boldsymbol{A}^{-1}\boldsymbol{U}$ is typically much easier to invert than the larger matrix $\boldsymbol{A} + \boldsymbol{U}\boldsymbol{V}^\intercal$.

(c) Show the binomial inversion formula

$$(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{B}\boldsymbol{V}^\intercal)^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}\boldsymbol{B}^{-1}(\boldsymbol{B}^{-1} + \boldsymbol{V}^\intercal\boldsymbol{A}^{-1}\boldsymbol{U})^{-1}\boldsymbol{B}\boldsymbol{V}^\intercal\boldsymbol{A}^{-1},$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are nonsingular.

(d) Show the identity

$$\det(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{V}^\intercal) = \det(\boldsymbol{A})\det(\boldsymbol{I}_m + \boldsymbol{V}^\intercal\boldsymbol{A}^{-1}\boldsymbol{U}).$$

This formula is useful for evaluating the density of a multivariate normal with covariance matrix $\boldsymbol{A} + \boldsymbol{U}\boldsymbol{U}^\intercal$.