1. Using your class notes, prepare a 1-2 paragraph summary of what we talked about in class in the last week. I do not want just a bulleted list of topics, I want you to use complete sentences and establish context (Why is what we have learned relevant? How does it connect with other things you have learned here or in other classes?). The more insight you give, the better.

2. Let \( X \) be a convex set in \( \mathbb{R}^N \), and let \( A \) be an \( M \times N \) matrix. Let 
\[
Y = \{ y \in \mathbb{R}^M : y = Ax \text{ for some } x \in X \}.
\]
Show that \( Y \) is convex.

3. Find an explicit equation for the finding the closest point in a hyperplane. For fixed \( a \in \mathbb{R}^N \) and \( b \in \mathbb{R} \), let \( H \) be the hyperplane
\[
H = \{ x \in \mathbb{R}^N : \langle x, a \rangle = b \}.
\]
We want a closed form solution to the problem
\[
\min_{x \in H} \| x_0 - x \|_2,
\]
where \( x_0 \) is a given point. Give a formula (depending on \( x_0, a, b \)) for both the minimizer \( \hat{x} \) for the problem above, and the distance \( \| x_0 - \hat{x} \|_2 \).

4. Let \( X_0 \) be an \( N \times N \) symmetric matrix with eigenvalue decomposition
\[
X_0 = V \Lambda V^T = \sum_{n=1}^{N} \lambda_n v_n v_n^T.
\]
Show that the closest symmetric positive semidefinite matrix to \( X_0 \) can be calculated simply by truncating the terms with negative eigenvalues in the expression above:
\[
\arg \min_{X \in S_+^N} \| X_0 - X \|_F = \sum_{n=1}^{N} \max(\lambda_n, 0) v_n v_n^T =: \hat{X}.
\]
Do this by establishing the “obtuseness” property for \( \hat{X} \),
\[
\langle X - \hat{X}, X_0 - \hat{X} \rangle \leq 0 \text{ for all } X \in S_+^N.
\]
(Recall that the inner product between two matrices is
\[
\langle X, Y \rangle = \sum_{n, \ell} X_{n, \ell} Y_{n, \ell} = \text{trace}(Y^T X).
\]

It is easy to see that \( \langle X, X \rangle = \|X\|_F^2 = \sum_{n, \ell} X_{n, \ell}^2 \). It is always useful to remember that if one of the matrices in the inner product is rank-1, then you can re-write the inner product as a quadratic form; for any \( v \in \mathbb{R}^N \), we have \( \langle X, vv^T \rangle = v^T X v \).)

5. Let \( \mathcal{C} \) and \( \mathcal{D} \) be closed convex sets in \( \mathbb{R}^N \), with \( \mathcal{C} \subset \mathcal{D} \).

(a) Does \( P_{\mathcal{C}}(P_{\mathcal{D}}(x_0)) = P_{\mathcal{C}}(x_0) \)? Prove that it is true or provide a counter example.

(b) Does your answer change if \( \mathcal{D} \) is a subspace?

(c) Given an arbitrary \( N \times N \) matrix \( X_0 \), how do you compute the closest symmetric matrix to \( X_0 \)?

(d) Given an arbitrary \( N \times N \) matrix \( X_0 \), how do you compute the closest symmetric positive semidefinite matrix to \( X_0 \)?

6. Consider the following graph:

The nodes correspond to Gaussian random variables and the edges between the nodes describe their relationship. Specifically, from a set of iid Gaussian random variables \( E_n \sim \text{Normal}(0, 1) \), we generate
\[
X_1 = E_1 \\
X_2 = a_{12}X_1 + E_2 \\
X_3 = a_{13}X_1 + E_3 \\
X_4 = a_{34}X_3 + E_4 \\
X_5 = a_{35}X_3 + E_5.
\]

Let \( X = [X_1 \ X_2 \ X_3 \ X_4 \ X_5]^T \) be the random vector consisting of the collection of these random variables. We can generate \( K \) independent realizations of \( X \) in MATLAB using
\[
X = \text{zeros}(5, K); \\
\text{for} \; \text{kk} = 1:K \\
E = \text{randn}(5, 1); \\
X(1,kk) = E(1); \\
X(2,kk) = a12*X(1,kk) + E(2); \\
\]
\begin{align*}
X(3,kk) &= a_{13}X(1,kk) + E(3); \\
X(4,kk) &= a_{34}X(3,kk) + E(4); \\
X(5,kk) &= a_{35}X(3,kk) + E(5);
\end{align*}

end

None of these random variables is independent of the others. That is, the covariance matrix $R = \mathbb{E}[XX^T]$ is non-zero everywhere (assuming, of course, that the specified $a_{i,j} \neq 0$). But it should be clear that pairs of nodes without edges between them are conditionally independent (given the values at the other nodes). So, for example:

\[ X_2 \mid (X_1, X_4, X_5) \text{ is independent of } X_3 \mid (X_1, X_4, X_5) \]

(And in fact in this case, $X_2$ and $X_3$ are independent given just the values of $X_1$.) It is a fact that if two entries $X_i, X_j$ in a Gaussian random vector are conditionally independent, then the corresponding entries in the inverse covariance matrix are zero; $S_{i,j} = S_{j,i} = 0$, where $S = R^{-1}$.

For this problem, we will fix the values $a_{12} = 2$, $a_{13} = -1$, $a_{34} = 1/2$, $a_{35} = -0.1$.

(a) Compute by hand the covariance matrix $R$. Verify (using MATLAB) that its inverse is zero in the appropriate places.

(b) In MATLAB, generate $K = 1000$ realizations of $X$. Compute the standard sample covariance

\[ \hat{R} = \frac{1}{K} \sum_{k=1}^{K} XX^T. \]

Describe how close it is to the true covariance matrix both quantitatively (using a reasonable metric of your choosing) and qualitatively.

(c) Now suppose that we know the structure of the graph above, but not the values of $a_{i,j}$. Recall (from our first set of notes) that the inverse of the sample covariance is the solution to the convex program

\[ \min_{X \in S_+^5} - \log \det X + \text{trace}(X \hat{R}). \]

We can exploit our knowledge of the graph structure above to help us get a better estimate of the covariance matrix from these same $K = 1000$ samples. Let $\mathcal{I} = \{(i,j) : \text{there is no edge between nodes } i \text{ and } j\}$. Now estimate the covariance matrix by using CVX to solve

\[ \min_X - \log \det X + \text{trace}(X \hat{R}), \text{ subject to } X \in S_+^5, \ X_{i,j} = 0, \ (i,j) \in \mathcal{I}. \]

and taking the inverse of your answer. Describe how close it is to the true covariance matrix both quantitatively and qualitatively.

(CVX handles the semidefinite constraint very naturally; see the section on Set Membership in \url{http://web.cvxr.com/cvx/doc/}. Also, the log det function is built into CVX: \url{http://web.cvxr.com/cvx/doc/funcref.html#built-in-functions}.)

(d) \textit{Optional}: Show that if $X_i, X_j$ are conditionally independent, then $R_{i,j}^{-1} = 0$. You can do this by combining the expression for the conditional covariance matrix for a Gaussian random vector and the Schur complement for $R$ (see Appendix C.4 of Boyd).