Linear Algebra and Applications

Benjamin Recht

February 28, 2005

The Gameplan

- Basics of linear algebra

- matrices as operators

- matrices as data

- matrices as costs/constraints

Basics of linear algebra

A vector is a collection of real numbers arranged in an array. Vectors can be multiplied by real numbers and added to one another.

Lowercase letters like $\mathbf{x}, \mathbf{y}, \mathbf{z}$ will denote arrays of size $n \times 1$. The set of all $n \times 1$ vectors is denoted $\mathbb{R}^n$. Capital letters like $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ will denote $m \times n$ dimensional arrays and are called *matrices*. The set of all $m \times n$ matrices is denoted $\mathbb{R}^{m \times n}$.

The entries of vectors and matrices are given by non-boldfaced letters. For example, the element in the ith row and jth column of the matrix $\mathbf{A}$ is A_{ij}.

Bases

- If $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are vectors, a linear combination is a sum $\sum_{k=1}^{N} a_k \mathbf{x}_k$ which is also a vector.

- A set of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in linearly independent if $\sum_{k=1}^{N} a_k \mathbf{x}_k = 0$ only when $a_k = 0$ for all k.

- A basis is a linearly independent set of n-vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ such that any n-vector $\mathbf{v}$ can be written as a linear combination of the $\mathbf{e}_k$. That is, $\mathbf{v} = \sum_{k=1}^{n} a_k \mathbf{e}_k$ for some a_k.

Matrices as Operators (1)

- *linearity*: If $\mathbf{x}, \mathbf{y}$ are vectors then $ax + by$ is a vector for any scalars a and b.

- *linearity (2)*: $f : \mathbb{R}^m \to \mathbb{R}^n$ is *linear* if $f(a\mathbf{x} + b\mathbf{y}) = af(\mathbf{x}) + bf(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ and scalars a, b.

- *fact*: If f is linear, then there is an $n \times m$ matrix $\mathbf{A}$ such that $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$.

- $\mathbf{A}$ $n \times m$, $\mathbf{B}$ $m \times l$, then $\mathbf{AB}$ is $n \times l$.

Matrices as Operators (2)

- Let $\mathbb{1}$ denote the map where $\mathbb{1}x = x$ for all x.

- $\mathbf{A}$ $n \times n$, if there exists a matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbb{1}$ then $\mathbf{A}^{-1}$ is called the *inverse* of $\mathbf{A}$.

- *FACT*: $\mathbf{A}$ is invertible (i.e., $\mathbf{A}$ has an inverse) if and only if the columns of $\mathbf{A}$ are linearly independent (and hence form a basis).

- $\mathbf{A}^\top$ is the *transpose* of $\mathbf{A}$. If A_{ij} is the entry in the ith row and jth column of $\mathbf{A}$, A_{ji} is the entry in the ith row and jth column of $\mathbf{A}^\top$. A matrix is symmetric if $\mathbf{A}^\top = \mathbf{A}$.

Matrices as operators (3)

- If $\mathbf{x}$ and $\mathbf{y}$ are vectors, $\mathbf{y}^\top\mathbf{x}$ is 1×1, a scalar. This is the *inner product* of $\mathbf{x}$ and $\mathbf{y}$.

- If $\mathbf{y}^\top\mathbf{x} = 0$ then $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal*. If furthermore $\mathbf{x}^\top\mathbf{x} = \mathbf{y}^\top\mathbf{x} = 1$ then the vectors are *orthonormal*.

- $\mathbf{x}\mathbf{y}^\top$ is $n \times n$. This is the *outer product* of $\mathbf{x}$ and $\mathbf{y}$

Matrices and Systems

An $n \times n$ matrix $\mathbf{A}$ can map n-vectors over time. Continuous system:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}(t)$$

Discrete time system:

$$\mathbf{x}[n] = \mathbf{A}\mathbf{x}[n-1]$$

Continuous Time Solution

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}(t) \qquad \text{ANSATZ: } \mathbf{x}(t) = \exp(\mathbf{A}t)\mathbf{x}(0)$$

Define:

$$\exp(\mathbf{A}t) = \sum_{k=0}^{\infty} \frac{1}{k!}(\mathbf{A}t)^k = \mathbb{1} + \mathbf{A}t + \frac{1}{2}\mathbf{A}^2t^2 + \frac{1}{6}\mathbf{A}^3t^3 + \ldots$$

taking d/dt gives

$$\frac{d}{dt}\exp(\mathbf{A}t) = \sum_{k=1}^{\infty} \frac{k}{k!}\mathbf{A}^k t^{k-1} = \mathbf{A}\sum_{k=1}^{\infty} \frac{1}{(k-1)!}\mathbf{A}^{k-1}t^{k-1} = \mathbf{A}\exp(\mathbf{A}t)$$

proving

$$\frac{d}{dt}\exp(\mathbf{A}t)x(0) = \mathbf{A}\exp(\mathbf{A}t)x(0)$$

Properites of the exponential map

- If $\mathbf{S}$ is invertible, $\exp(\mathbf{SAS}^{-1}) = \mathbf{S}\exp(\mathbf{A})\mathbf{S}^{-1}$.

- If $\mathbf{D}$ is diagonal, $\mathbf{E} = \exp(\mathbf{D})$ is diagonal and $E_{jj} = \exp(D_{jj})$

Discrete Time Solution

$$x[n] = A^n x[0]$$

Here the proof is immediate.

Analysis

What can we say about linear systems without simulation?

- Does the system oscillate?
- Does the system converge to zero?
- Does the system diverge to infinity?

Eigenvalues

If $\mathbf{A}$ is an $n \times n$ matrix, λ is an *eigenvalue* of $\mathbf{A}$ if $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ for some $\mathbf{v} \neq \mathbf{0}$. $\mathbf{v}$ is an *eigenvector*.

FACT: If $\mathbf{v}_1$ and $\mathbf{v}_2$ are both eigenvectors of $\mathbf{A}$ with eigenvalues $\lambda_1 \neq \lambda_2$, then $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

Proof By contradiction, assume there exist nonzero a and b such that

$$a\mathbf{v}_1 + b\mathbf{v}_2 = \mathbf{0} \implies A(a\mathbf{v}_1 + b\mathbf{v}_2) = \mathbf{0} \implies a\lambda_1\mathbf{v}_1 + b\lambda_2\mathbf{v}_2 = \mathbf{0}$$

Multiply the first equation by λ_1 and subtract to find

$$b(\lambda_2 - \lambda_1)\mathbf{v}_2 = \mathbf{0}$$

which is a contradiction.

■

FACT: If $\mathbf{A}$ has n distinct eigenvalues then it's eigenvectors are linearly independent.

FACT: If $\mathbf{S} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ then $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is a diagonal matrix.

Proof

$$\begin{aligned}\mathbf{S}^{-1}\mathbf{A}\mathbf{S} &= \mathbf{S}^{-1}[\lambda_1\mathbf{x}_1, \ldots, \lambda_n\mathbf{x}_n] \\ &= \mathbf{S}^{-1}[\mathbf{x}_1, \ldots, \mathbf{x}_n]\mathbf{D} \\ &= \mathbf{S}^{-1}\mathbf{S}\mathbf{D} = \mathbf{D}\end{aligned}$$

■

FACT: If $\mathbf{A}$ is symmetric, $\mathbf{A}\mathbf{x} = \lambda_1 \mathbf{x}$ and $\mathbf{A}\mathbf{y} = \lambda_2 \mathbf{y}$ then $\lambda_1 \neq \lambda_2 \implies \mathbf{y}^\top \mathbf{x} = 0$.

Proof

$$\begin{aligned} \mathbf{y}^\top \mathbf{A}\mathbf{x} = \mathbf{y}^\top (\mathbf{A}\mathbf{x}) &= \lambda_1 \mathbf{y}^\top \mathbf{x} \\ = (\mathbf{A}\mathbf{y})^\top \mathbf{x} &= \lambda_2 \mathbf{y}^\top \mathbf{x} \end{aligned}$$

Since $\lambda_1 \neq \lambda_2$, we have $\mathbf{y}^\top \mathbf{x} = 0$ ∎

Corollary: The eigenvectors of a symmetric matrix $\mathbf{A}$ may be chosen to be orthonormal. If $\mathbf{S} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ then $\mathbf{S}^\top \mathbf{S} = \mathbb{1}$ and $\mathbf{S}^\top \mathbf{A}\mathbf{S}$ is a diagonal matrix.

Stability: Continuous Time

Suppose $\mathbf{A}$ has n distinct eigenvalues or is symmetric. Then

$$\mathbf{x}(t) = \exp(\mathbf{A}t)x(0) = \mathbf{S}^{-1}\exp(\mathbf{D}t)\mathbf{S}x(0)$$

- If $\mathrm{Re}(\lambda_n) < 0$ for all n, $\lim_{t\to\infty}\mathbf{x}(t) = \mathbf{0}$. The system is *stable*

- If $\mathrm{Re}(\lambda_n) \leq 0$ for all n, $\|\mathbf{x}(t)\| < \infty$ for all t. The system is *oscillating*.

- If $\mathrm{Re}(\lambda_n) > 0$ for any n, then $\lim_{t\to\infty}\mathbf{x}(t) = \infty$. The system is *unstable*.

Stability: Discrete Time

Suppose $\mathbf{A}$ has n distinct eigenvalues or is symmetric. Then

$$\mathbf{x}[k] = \mathbf{A}^k x[0] = \mathbf{S}^{-1}\mathbf{D}^k\mathbf{S}x(0)$$

- If $|\lambda_n| < 1$ for all n, $\lim_{k\to\infty} \mathbf{x}[k] = \mathbf{0}$. The system is *stable*
- If $|\lambda_n| \leq 1$ for all n, $\|\mathbf{x}[k]\| < \infty$ for all k. The system is *oscillating*.
- If $|\lambda_n| > 1$ for any n, then $\lim_{k\to\infty} \mathbf{x}[k] = \infty$. The system is *unstable*.

Random Vectors

Suppose we are observing a process described by a list of d numbers

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}$$

If each x_i are random variables $\mathbf{x}$ is a random vector.

The joint probability distribution is given by $p(\mathbf{x})$. We have

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\mathbf{x}) dx_1 \ldots dx_d = 1$$

Moments and Expectations

- The expected value of a matrix valued function $\mathbf{A}(\mathbf{x})$ is a matrix $\mathbb{E}[\mathbf{A}]$ with entries $\int A_{jk}(\mathbf{x})p(\mathbf{x})d\mathbf{x}$.

- The *mean*: $\bar{\mathbf{x}} = \mathbb{E}[\mathbf{x}]$.

- The *correlation*: $\mathbf{R}_{\mathbf{x}} = \mathbb{E}[\mathbf{x}\mathbf{x}^\top]$.

- The *covariance*: $\Lambda_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}} - \bar{\mathbf{x}}\bar{\mathbf{x}}^\top$

Positive Semidefinite Matrices

A matrix $\mathbf{Q}$ is *positive semidefinite* (psd) if $\mathbf{Q} = \mathbf{Q}^\top$ and for all $\mathbf{x}$, $\mathbf{x}^\top \mathbf{Q} \mathbf{x} \geq 0$. This is denoted $\mathbf{Q} \succeq 0$

- If $\mathbf{P} \succeq 0$ and $\mathbf{Q} \succeq 0$ and $a > 0$, then $a\mathbf{Q} \succeq 0$ and $\mathbf{Q} + \mathbf{P} \succeq 0$.

- If $\mathbf{Q} \succeq 0$, then $\mathbf{Q}$ is diagonalizable and has only nonnegative eigenvalues. Moreover, the eigenvectors can be chosen to be orthonormal.

- The outer product $\mathbf{x}\mathbf{x}^\top$ is psd. The correlation and covariances matrices of a random vector are psd.

Matrices as Data

N data points $\mathbf{x}_1, \ldots \mathbf{x}_N$ each consisting of a list of d numbers $\mathbf{x}_n = \{x_{1n}, \ldots, x_{dn}\}$

- Images (e.g., 640x480 pixels)
- Audio (e.g., samples)
- Diagnostics (e.g., lab results)

The data matrix is defined to be $\mathbf{X}$ where $X_{ij} = x_{ij}$. It is $d \times N$.

Empirical Statistics

- *mean* $\bar{\mathbf{x}}_N = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i$
- *zero-mean data matrix* $\widehat{X}_{N,ij} = x_{ij} - \bar{x}_{N,i}$
- *covariance* $\Lambda_N = \frac{1}{N}\widehat{\mathbf{X}}_N\widehat{\mathbf{X}}_N^\top$. $\Lambda_N \succeq 0$
- *gram matrix* $\mathbf{K}_N = \widehat{\mathbf{X}}_N^\top\widehat{\mathbf{X}}_N$. $\mathbf{K}_N \succeq 0$

Multivariate Gaussians

$$p(\mathbf{x}) = \frac{1}{\sqrt{|2\pi\Lambda|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^\top \Lambda (\mathbf{x}-\bar{\mathbf{x}})\right)$$

The mean of this random vector is $\bar{\mathbf{x}}$. The covariance is Λ.

Since $\Lambda \succeq 0$, there is a matrix $\mathbf{C}$ such that $\Lambda = \mathbf{C}^\top \Delta \mathbf{C}$ with $\mathbf{C}^\top \mathbf{C} = \mathbb{1}$ and Δ is a diagonal matrix. The random vector $\mathbf{y} = \mathbf{C}(\mathbf{x}-\bar{\mathbf{x}})$ has zero mean and covariance Δ. That means the components of $\mathbf{y}$ are independent.

Marginal and Conditional Moments

- Let $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$ with $\mathbf{v}$ constant. Then $\bar{\mathbf{y}} = \mathbf{A}\bar{\mathbf{x}} + \mathbf{v}$ and $\Lambda_y = \mathbf{A}\Lambda_x\mathbf{A}^\top$

- Let $z = [\mathbf{x}, \mathbf{y}]$ be gaussian. Then

$$\bar{x} = \begin{bmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{y}} \end{bmatrix} \quad \text{and} \quad \Lambda_z \equiv \begin{bmatrix} \Lambda_x & \Lambda_{xy} \\ \Lambda_{yx} & \Lambda_y \end{bmatrix}$$

  $\Lambda_{xy} = \Lambda_{yx}^\top$

- $p(\mathbf{x}|\mathbf{y})$ is a gaussian with mean $\bar{x}+\Lambda_{xy}\Lambda_y^{-1}(\mathbf{y}-\bar{\mathbf{y}})$ and variance $\Lambda_x - \Lambda_{xy}\Lambda_y^{-1}\Lambda_{xy}^\top$.

PCA

Given a zero-mean random vector $\mathbf{x}$, let us suppose that we want to represent $\mathbf{x}$ as a $\sum_i a_i \mathbf{y}_i$ with the $\mathbf{y}_i$ uncorrelated. Then the best solution is to have $\mathbf{y}_i = \mathbf{C}_i$, $a_i = \Delta_{ii}$.

When we only have finitely many examples, Λ_N is the best estimate of the actual covariance. Given a matrix of data $\mathbf{X}$, $\Lambda_N = \mathbf{C}_N \Delta_N \mathbf{C}_N^\top$. So we can use $\mathbf{y}_i = \mathbf{C}_{N,i}$ and $a_i = \Delta_{N,ii}$.

If we only have a small number of data points as compared to dimensions, diagonalizing Λ_N can be very computer intensive. The Singular Value Decomposition makes this tractable.

Singular Value Decomposition

If $\mathbf{A}$ is a matrix of size $m \times n$ then there exists orthogonal matrices $\mathbf{V}$ $(m \times m)$ and $\mathbf{W}$ $(n \times n)$ such that

$$\mathbf{V}^\top \mathbf{A} \mathbf{W} = \text{diag}(\sigma_1, \ldots, \sigma_p)$$

with $\sigma_1 \geq \ldots \geq \sigma_p \geq 0$, $p = \min(m, n)$.

Proof Without loss of generality, assume $m \leq n$. Since $\mathbf{A}^\top \mathbf{A} \succeq 0$, the eigenvalues of $\mathbf{A}^\top \mathbf{A}$ are equal to $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \sigma_n^2 \geq 0$ for some $\sigma_k \geq 0$. Let r be the largest number for which $\sigma_r > 0$.

Let $\mathbf{x}_k$ be norm 1 eigenvectors of $\mathbf{A}^\top\mathbf{A}$ corresponding to σ_k^2 for $k = 1, \ldots, r$. Let $\mathbf{y}_k = \mathbf{A}\mathbf{x}_k/\sigma_k$. Since $\|\mathbf{A}\mathbf{x}_k\|^2 = \mathbf{x}_k^\top\mathbf{A}^\top\mathbf{A}\mathbf{x}_k = \sigma_k^2$, $\mathbf{y}_k$ are norm 1. Furthermore, $\mathbf{y}_j^\top\mathbf{y}_k = \frac{1}{\sigma_j\sigma_k}\mathbf{x}_j^\top\mathbf{A}^\top\mathbf{A}x_k = 0$ when $k \neq j$ so the $\mathbf{y}_k$ are orthonormal.

Completing $\mathbf{y}_k$ to an orthonormal basis for $\mathbb{R}^n$ gives matrices $\mathbf{W} \equiv [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ and $\mathbf{V} \equiv [\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{V}_2]$. It is easy algebra to check that $\mathbf{V}^\top\mathbf{A}\mathbf{W}$ has the desired form (See Problem 2). ■

SVD(A,0)

When $\mathbf{A}$ is a matrix of $m \times n$ with $m > n$, we would rather compute the eigenvalues of $\mathbf{A}^\top\mathbf{A}$ than of $\mathbf{A}\mathbf{A}^\top$. Furthermore, we need only compute the first n columns of $\mathbf{V}$ for $\mathbf{A} = \mathbf{V}\mathbf{S}\mathbf{W}^\top$ to hold.

The matlab command:

```
[V,S,W]=svd(A,0)
```

performs this computation efficiently.

Matrices as cost/constraints

We will frequently encounter cost functions and constraints defined by matrices:

- *linear equalities*: $\mathbf{A}\mathbf{x} = \mathbf{b}$
- *linear inequalities*: $\mathbf{A}\mathbf{x} \leq \mathbf{b}$
- *linear cost*: $c(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$
- *quadratic cost*: $c(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$
- *least squares*: $c(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$

Linear Constraints

Note that linear equalities and linear inequalities are interchangeable by adding constraints or variables:

$$\mathbf{Ax} = \mathbf{b} \iff \mathbf{Ax} \leq \mathbf{b} \quad \text{and} \quad \mathbf{Ax} \geq \mathbf{b}$$

$$\mathbf{Ax} \leq b \iff \mathbf{Ax} = \mathbf{b} + \mathbf{s} \quad \text{and} \quad \mathbf{s} \geq 0$$

Such $\mathbf{s}$ are called *slack variables*

Unconstrained Quadratic Programming

$$\min_{\mathbf{x}} \mathbf{x}^\top \mathbf{A}\mathbf{x} = \begin{cases} 0 & A \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\min_{\mathbf{x}} \mathbf{x}^\top \mathbf{A}\mathbf{x} - 2\mathbf{b}^\top \mathbf{x} + c$$

Differentiate with respect to $\mathbf{x}$ to find that at the optimum

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

If $\mathbf{A}$ is invertible then the minimum is $-\mathbf{b}^\top \mathbf{A}^{-1}\mathbf{b} + c$

Schur Complements

Let

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}$$

The Schur complement of $\mathbf{C}$ in $\mathbf{A}$ is given by

$$(\mathbf{M}|\mathbf{A}) = \mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}$$

Similarly

$$(\mathbf{M}|\mathbf{C}) = \mathbf{A} - \mathbf{B}\mathbf{C}^{-1} \mathbf{B}^\top$$

Facts about Schur complements:

$$\mathbf{M} \succeq 0 \iff \mathbf{C} \succeq 0 \quad \text{and} \quad (\mathbf{M}|\mathbf{C}) \succeq 0$$

$$\mathbf{M}^{-1} = \begin{bmatrix} (\mathbf{M}|\mathbf{C})^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{M}|\mathbf{A})^{-1} \\ -\mathbf{C}^{-1}\mathbf{B}^{\top}(\mathbf{M}|\mathbf{C})^{-1} & (\mathbf{M}|\mathbf{A})^{-1} \end{bmatrix}$$

More Quadratic Programming

For the quadratic minimization

$$\min_{\mathbf{x}_2} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}^\top \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} - 2 \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}^\top \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$$

$$\mathbf{x}_2^* = \mathbf{C}^{-1}(\mathbf{b}_2 - \mathbf{B}\mathbf{x}_1)$$

Plug that back into the cost function:

$$\mathbf{x}_1^\top (\mathbf{M}|\mathbf{C})\mathbf{x}_1 - 2(\mathbf{b}_1 - \mathbf{B}\mathbf{C}^{-1}\mathbf{b}_2)^\top \mathbf{x}_1$$

Matrix Inversion Lemma

$$(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top)^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{B}^\top\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{A}^{-1}$$

To check this, apply the partitioned matrix formula twice and set the first blocks equal to each other.

Standard form: $\mathbf{C} \to -\mathbf{C}^{-1}$

$$(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{B}^\top)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{B}^\top\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{A}^{-1}$$

Problem1: Emergence of thermodynamics

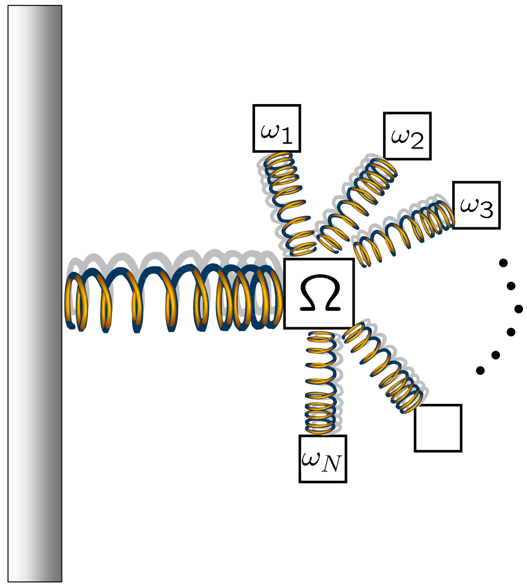
The following ODE describes the time evolution of a set of coupled masses and springs called the Caldeira-Leggett (CL) model

$$\frac{dx_k}{dt} = p_k/m_k \quad \text{for } k = 0, \ldots, N$$

$$\frac{dp_0}{dt} = -m_0\Omega^2 x_0 + \sum_{k=1}^{N} g_k(\omega_k x_k - g_k x_0/m_k)$$

$$\frac{dp_k}{dt} = -m_k\omega_k^2 x_k + g_k\omega_k x_0 \quad \text{for } k = 1, \ldots, N$$

x_k and p_k respectively denote the position and momenta of the kth spring (see figure).



- Let $N = 200$, $\Omega = 1$, $\gamma = 1$, $m_k = 1$, $\omega_k = 10k/N$, and $g_k = \sqrt{40\gamma/(N\pi)}$. Is the system stable, oscillatory, or unstable? How many oscillatory modes are there?

- Write a program to compute $x_0(t)$ with the initial condition $x_0(0) = 1$, $x_k(0) = 0$ for all $k = 1, \ldots, N$. Plot $x_0(t)$ from $t = 0$ to $t = 100$.

- Consider the system

$$\frac{dQ}{dt} = P/m_0$$
$$\frac{dP}{dt} = -m_0(\Omega^2 + \gamma^2)Q - 2\gamma P$$

with the same parameter settings as above. Is this system stable, oscillatory, or unstable? Analytically compute $Q(t)$ as a function of time. Plot $Q(t)$ for $t = 0$ to $t = 100$ with $Q(0) = 1$ and compare to the output of the CL model.

Problem2: Eigenfaces

Download the database of faces off the class website.

- Finish the proof of the singular value decomposition. That is, verify that

$$\mathbf{V}^\top \mathbf{A}\mathbf{W} = \text{diag}(\sigma_1, \ldots, \sigma_p)$$

- Compute the SVD of the data matrix. What do the principle components look like as images?

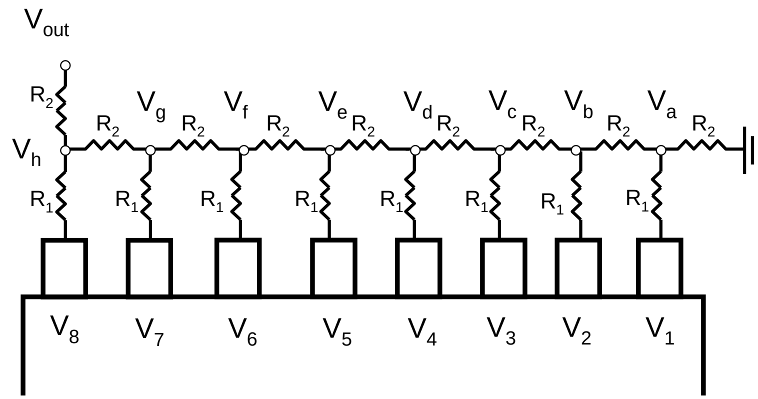
- Generate 4 a_1, a_2, a_3, a_4 numbers drawn independently from a gaussian and compute the image

  $$a_1\sigma_1\mathbf{V}_1 + a_2\sigma_2\mathbf{V}_2 + a_3\sigma_3\mathbf{V}_3 + a_4\sigma_4\mathbf{V}_4$$

  This is an *eigenface*

- Compute an *eigensomething* for a something of your choice.

Problem 3: Networks of Resistors



Recall from electronics that the voltage drop across a resistor (i.e., the difference of the voltages at either end) is equal to the current across the resistor times the voltage. Furthermore, remember that the sum of all currents into a node must equal zero. In equations that is:

$$gV = I, \qquad \sum_{I_i \in \mathcal{N}a} I_i = 0$$

where $g = 1/R$ is the conductance of a resistor.

- Write down these two conditions as matrix constraints on the resistor network. That is, find an 17×18 matrix $\mathbf{G}$ and an 8×17 matrix $\mathbf{K}$ such that

$$\mathbf{GV} = \mathbf{I} \quad \text{and} \quad \mathbf{KI} = \mathbf{0}$$

- If n is a number between 0 and 255, let $b_8 b_7 b_6 b_5 b_4 b_3 b_2 b_1$ be the binary expansion. If $R_1 = 20K$ and $R_2 = 10_K$, what is V_{out} when $V_i = b_i$ for $i = 1, \ldots 8$?