L. Vandenberghe
ECE133A 4/10/2021

Homework 1 solutions

1. *Exercise T1.17.* We are asked to write the T-vector

$$c = (1, 0, \ldots, 0, -(1+r)^{T-1})$$

as a linear combination of the $T-1$ vectors

$$l_t = (0, \ldots, 0, 1, -(1+r), 0, \ldots, 0), \quad t = 1, \ldots, T-1.$$

In the definition of l_t there are $t-1$ leading and $T-t-1$ trailing zeros. The element 1 is in position t. There is only one way to do this:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ -(1+r)^{T-1} \end{bmatrix} = \begin{bmatrix} 1 \\ -(1+r) \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} + (1+r) \begin{bmatrix} 0 \\ 1 \\ -(1+r) \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$
$$+ (1+r)^2 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -(1+r) \\ \vdots \\ 0 \\ 0 \end{bmatrix} + \cdots + (1+r)^{T-2} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ -(1+r) \end{bmatrix}.$$

In vector notation,

$$c = l_1 + (1+r)l_2 + (1+r)^2 l_3 + \cdots + (1+r)^{T-2} l_{T-1}.$$

The coefficients in the linear combination are

$$1, \quad 1+r, \quad (1+r)^2, \quad \ldots, \quad (1+r)^{T-1}.$$

The idea is that you extend the length of an initial loan by taking out a new loan each period to cover the amount that you owe. So after taking out a loan for \$1 in period 1, you take out a loan for \$$(1+r)$ in period 2, and end up owing \$$(1+r)^2$ in period 3. Then you take out a loan for \$$(1+r)^2$ in period 3, and end up owing \$$(1+r)^3$ in period 4, et cetera.

2. *Exercise T2.5.*

   (a) We note that

   $$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (1/2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1/2)\begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

   This is a linear combination with coefficients that add up to one, so we must have

   $$\psi(1,-1) = (1/2)\psi(1,0) + (1/2)\psi(1,-2) = 3/2.$$

   (b) The value of $\psi(2,-2)$ cannot be determined. For example, the two affine functions

   $$\psi(x_1,x_2) = 1 - (1/2)x_2, \qquad \psi(x_1,x_2) = 2x_1 - (1/2)x_2 - 1$$

   both satisfy $\psi(1,0) = 1$ and $\psi(1,-2) = 2$, but have a different value at $(2,-2)$. We can also see this from the general expression $\psi(x_1,x_2) = a_1x_1 + a_2x_2 + b$ of an affine function of two variables. From the given values of ψ at $(1,0)$ and $(1,-2)$,

   $$a_1 + b = 1, \qquad a_1 - 2a_2 + b = 2.$$

   Therefore $a_2 = (a_1 + b - 2)/2 = -1/2$ but a_1 and b are not uniquely defined.

3. *Exercise T3.25.*

   (a) The mean return of the portfolio is the average of the vector p:

   $$\begin{aligned} \mathbf{avg}(p) &= \mathbf{avg}(\theta r + (1-\theta)\mu^{\mathrm{rf}}\mathbf{1}) \\ &= \theta\mathbf{avg}(r) + (1-\theta)\mu^{\mathrm{rf}}\mathbf{avg}(\mathbf{1}) \\ &= \theta\mu + (1-\theta)\mu^{\mathrm{rf}}. \end{aligned}$$

   On the last line we use $\mathbf{avg}(r) = \mu$ and $\mathbf{avg}(\mathbf{1}) = 1$.
   The risk is the standard deviation of the vector p:

   $$\begin{aligned} \mathbf{std}(p) &= \frac{1}{\sqrt{T}}\|p - \mathbf{avg}(p)\mathbf{1}\| \\ &= \frac{1}{\sqrt{T}}\|\theta r + (1-\theta)\mu^{\mathrm{rf}}\mathbf{1} - (\theta\mu + (1-\theta)\mu^{\mathrm{rf}})\mathbf{1}\| \\ &= \frac{1}{\sqrt{T}}\|\theta(r - \mu\mathbf{1})\| \\ &= \frac{|\theta|}{\sqrt{T}}\|r - \mu\mathbf{1}\| \\ &= |\theta|\mathbf{std}(r) \\ &= |\theta|\sigma. \end{aligned}$$

   On line 2 we use the expression for $\mathbf{avg}(p)$ that we derived in part (a). The last step is the definition of $\sigma = \mathbf{std}(r)$.

(b) To achieve the target risk σ^{tar}, we need $|\theta| = \sigma^{\mathrm{tar}}/\sigma$, so there are two choices:

$$\theta = \sigma^{\mathrm{tar}}/\sigma, \qquad \theta = -\sigma^{\mathrm{tar}}/\sigma.$$

To choose the sign of θ we consider the portfolio return

$$\mathbf{avg}(p) = \theta\mu + (1-\theta)\mu^{\mathrm{rf}} = \mu^{\mathrm{rf}} + \theta(\mu - \mu^{\mathrm{rf}}).$$

To maximize this, for given $|\theta|$, we choose θ positive if $\mu > \mu^{\mathrm{rf}}$ and θ negative if $\mu < \mu^{\mathrm{rf}}$. This means we short the asset when its return μ is less than the risk-free return μ^{rf}.

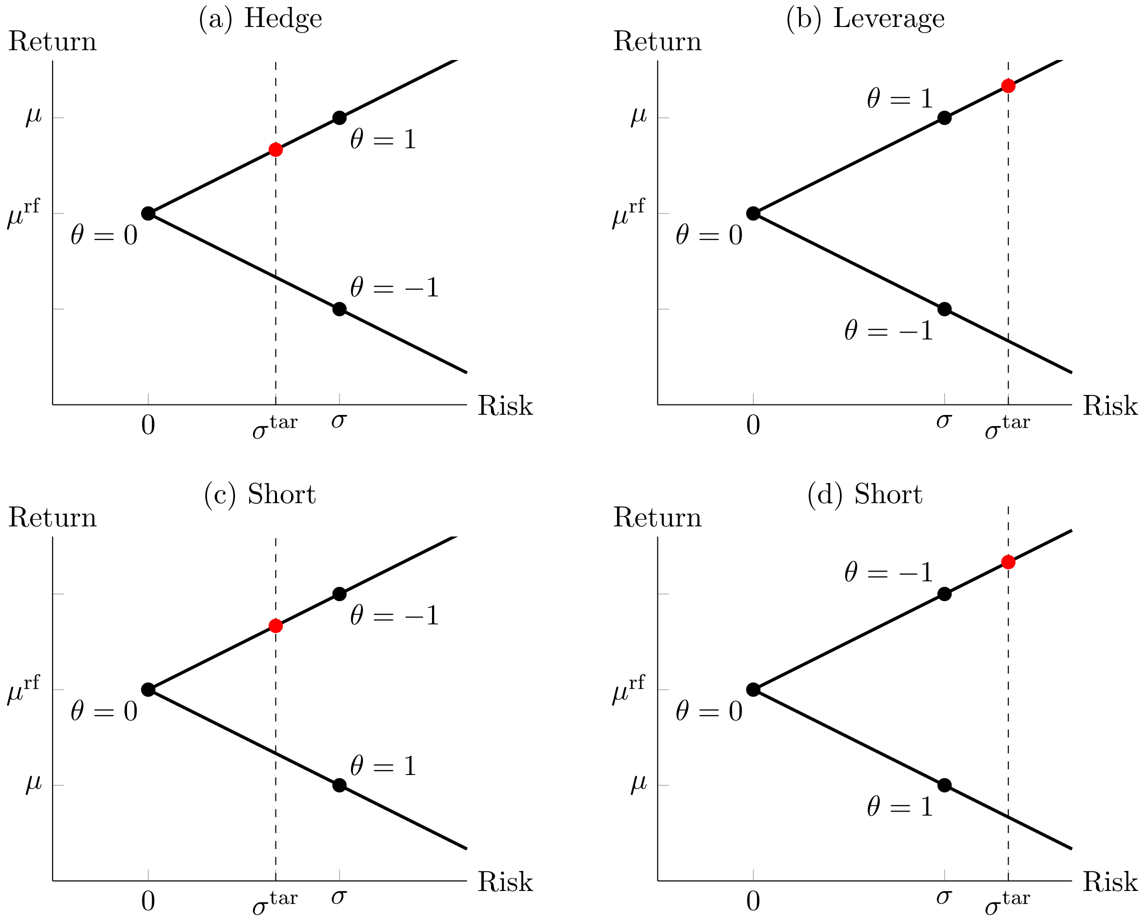
(c) We can distinguish the four cases shown in the figure on the next page. The solid lines show

$$\left[\begin{array}{c} \text{Risk} \\ \text{Return} \end{array}\right] = \left[\begin{array}{c} \mathbf{std}(p) \\ \mathbf{avg}(p) \end{array}\right] = \left[\begin{array}{c} |\theta|\sigma \\ \mu^{\mathrm{rf}} + \theta(\mu - \mu^{\mathrm{rf}}) \end{array}\right]$$

for all values of θ. The red dot shows the portfolio with the hightest return for the given target value of risk.

In case (a), the asset return is more than the risk-free return ($\mu > \mu^{\mathrm{rf}}$) and its risk is higher than the target risk ($\sigma > \sigma^{\mathrm{tar}}$). In this case we hedge ($0 < \theta < 1$). In case (b), the asset return is more than the risk-free return ($\mu > \mu^{\mathrm{rf}}$) and its risk is less than the target risk ($\sigma < \sigma^{\mathrm{tar}}/\sigma$). In this case we leverage ($\theta > 1$). In cases (c) and (d), the asset return is less than the risk-free return ($\mu < \mu^{\mathrm{f}}$). In this case we short ($\theta < 0$).

To summarize, we short the asset when its return is less than the risk-free return. We hedge when the asset return is more than the risk-free return and the asset risk is higher than the target risk. We leverage when the asset return is more than the risk-free return and the asset risk is less than the target risk.


(a) Hedge
Return
μ
μ^{rf}
$\theta = 0$
$\theta = 1$
$\theta = -1$
Risk
0
σ^{tar}
σ
(b) Leverage
Return
μ
μ^{rf}
$\theta = 0$
$\theta = 1$
$\theta = -1$
Risk
0
σ
σ^{tar}
(c) Short
Return
μ^{rf}
μ
$\theta = 0$
$\theta = -1$
$\theta = 1$
Risk
0
σ^{tar}
σ
(d) Short
Return
μ^{rf}
μ
$\theta = 0$
$\theta = -1$
$\theta = 1$
Risk
0
σ
σ^{tar}

4. *Exercise T4.3.* If $i \in G_1$, then x_i is closer to z_1 than to z_2:

$$\|x_i - z_2\|^2 - \|x_i - z_1\|^2 \geq 0.$$

If $i \in G_2$, then x_i is closer to z_2 than to z_1:

$$\|x_i - z_2\|^2 - \|x_i - z_2\|^2 \leq 0.$$

Expanding the norms on the left-hand side gives

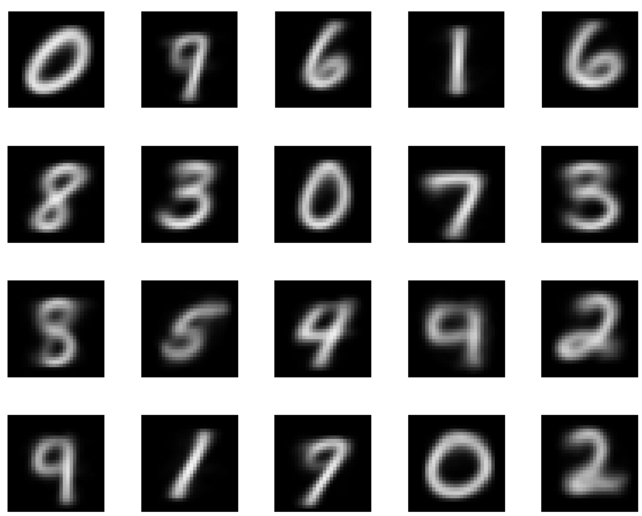
$$\begin{aligned} \|x_i - z_2\|^2 - \|x_i - z_1\|^2 &= \|x_i\|^2 - 2z_2^T x_i + \|z_2\|^2 - \|x_i\|^2 + 2z_1^T x_i - \|z_1\|^2 \\ &= -2(z_2 - z_1)^T x_i + \|z_2\|^2 - \|z_1\|^2. \end{aligned}$$

We see that if we take

$$w = -2(z_2 - z_1), \qquad v = \|z_2\|^2 - \|z_1\|^2,$$

then $w^T x_i + v \geq 0$ for $i \in G_1$ and $w^T x_i + v \leq 0$ for $i \in G_2$.

5. *Exercise A1.10.* A typical result is shown below.



It was computed using the following MATLAB code.

```
load mnist_train;
digits = digits(:,1:10000);
[n, N] = size(digits);
K = 20;
```

```
class = randi(K, 1, N);
Z = zeros(n,K);
D = zeros(K, N);
Jprev = NaN;
for iter = 1:100
    for i = 1:K
        % column i of Z is the i-th representative
        I = find(class == i);
        Z(:,i) = mean(digits(:,I), 2);
    end
    for i = 1:K
        % D(i,j) is the distance from example j to representative i
        D(i,:) = sqrt( sum( (digits - Z(:, i*ones(1,N))).^2) );
    end;
    % d(j) is the distance of example j to the nearest representative
    % class(j) is the index of the nearest representative
    [d, class] = min(D);
    J = (1/N) * norm(d)^2;
    if iter > 1
       if abs(J - Jprev) < 1e-5 * J, break; end;
       Jprev = J;
    end;
end;
for i=1:K
   subplot(4,5,i);
   imshow(reshape(Z(:,i), 28, 28));
end
```

In Julia we can solve the problem using the `kmeans` function from the companion, which is also available in the `VMLS` package.

```
using MAT, VMLS, LinearAlgebra, ImageView
f = matopen("mnist_train.mat")
labels = read(f,"labels")
labels = labels[1:10000]
digits = read(f,"digits")
digits = digits[:, 1:10000];
assignment, reps = kmeans(digits, 20);
X = [ reshape(hcat(reps[1:5]...), (28, 28*5));
      reshape(hcat(reps[6:10]...), (28, 28*5));
      reshape(hcat(reps[11:15]...), (28, 28*5));
      reshape(hcat(reps[16:20]...), (28, 28*5)) ];
imshow(X);
```